

Figure 1

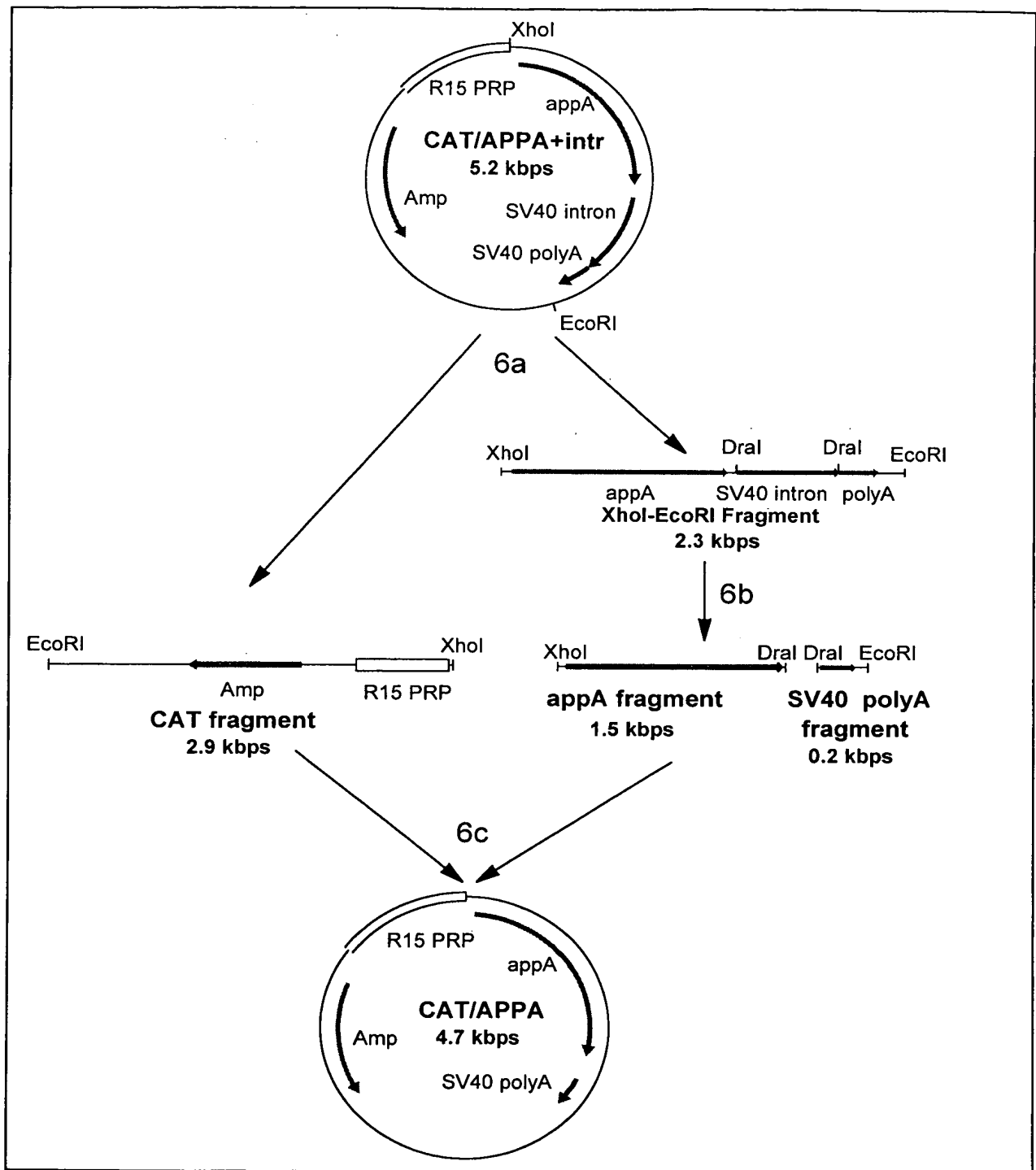


Figure 1A

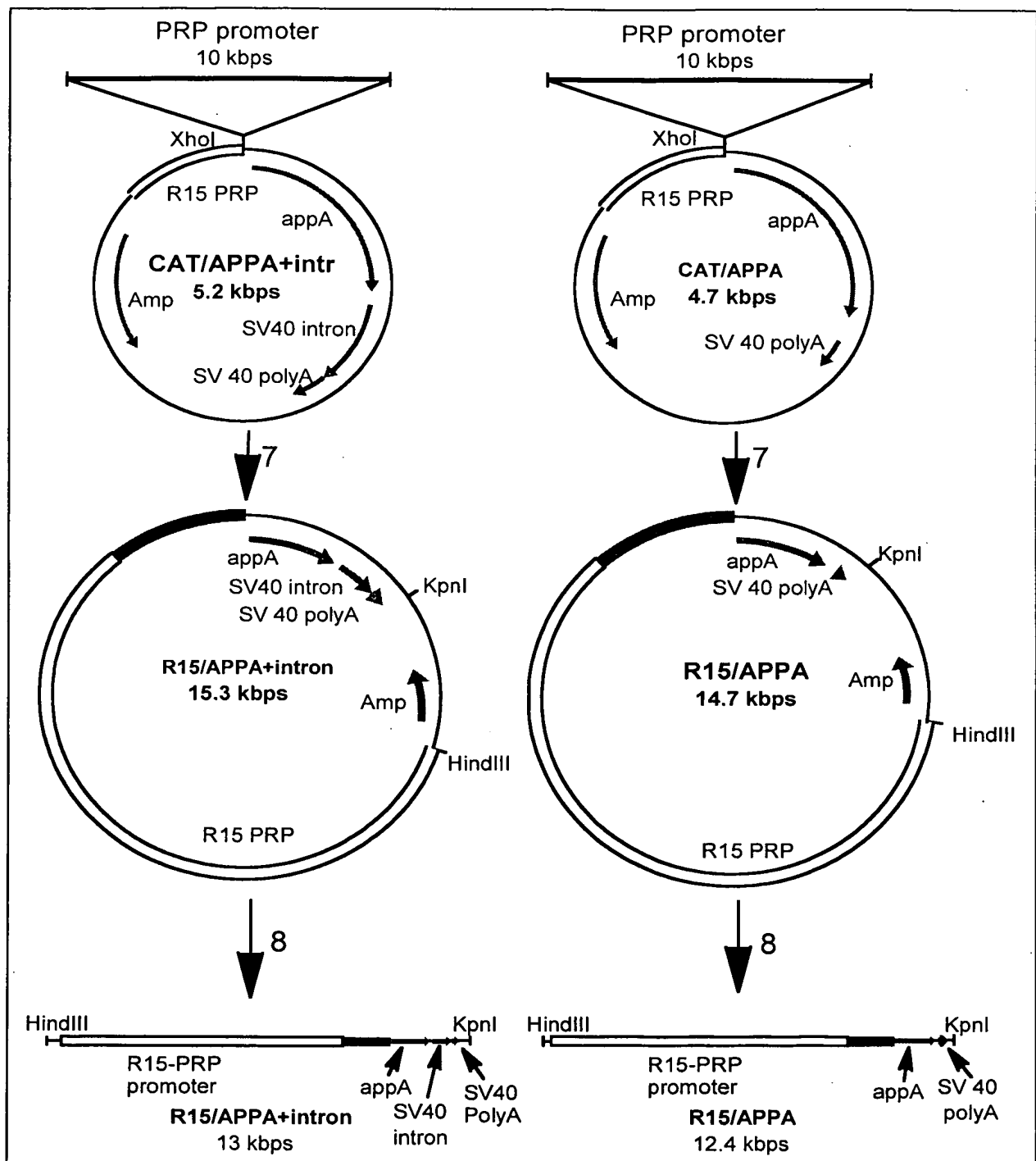


Figure 1B

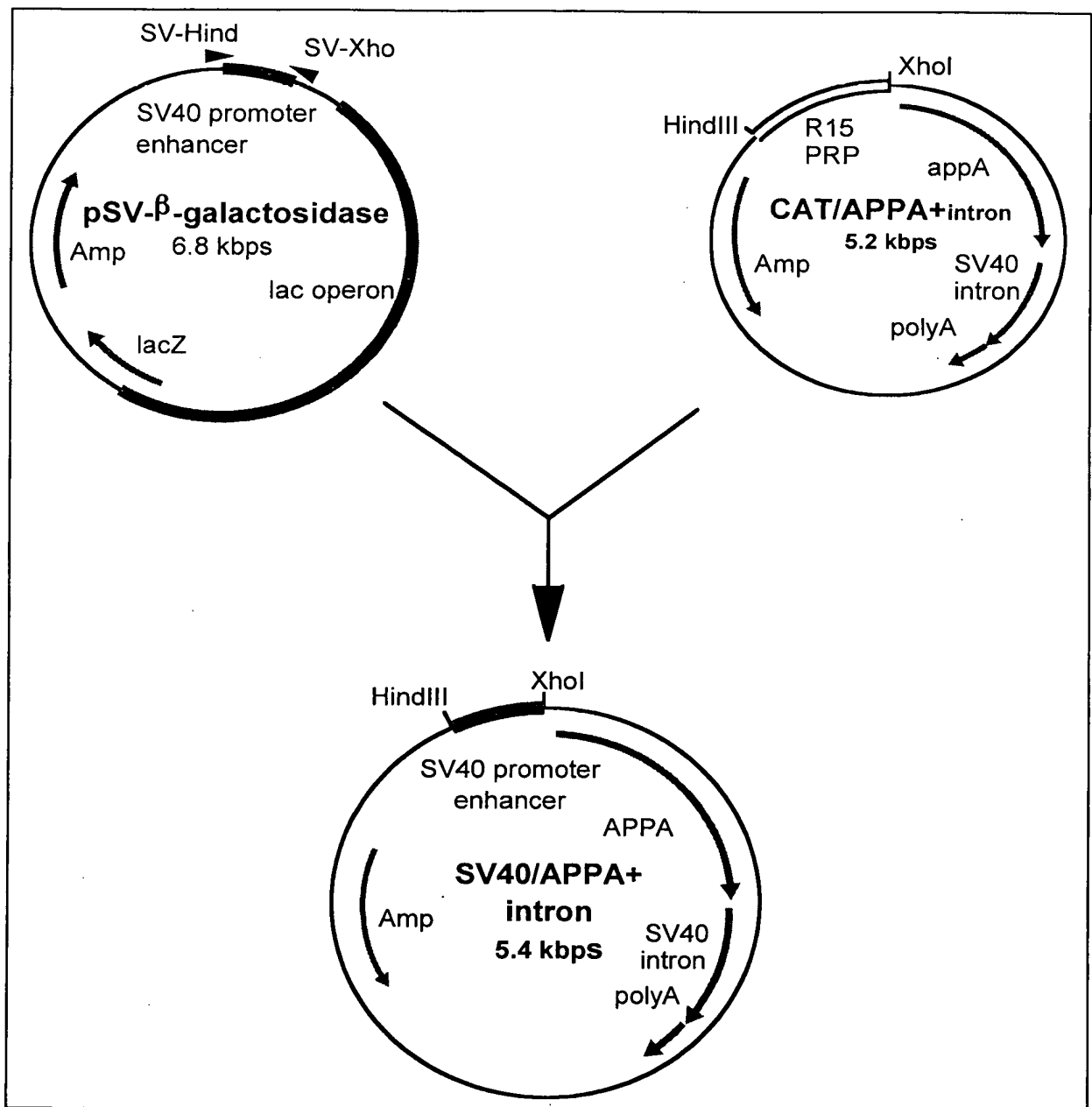


Figure 2

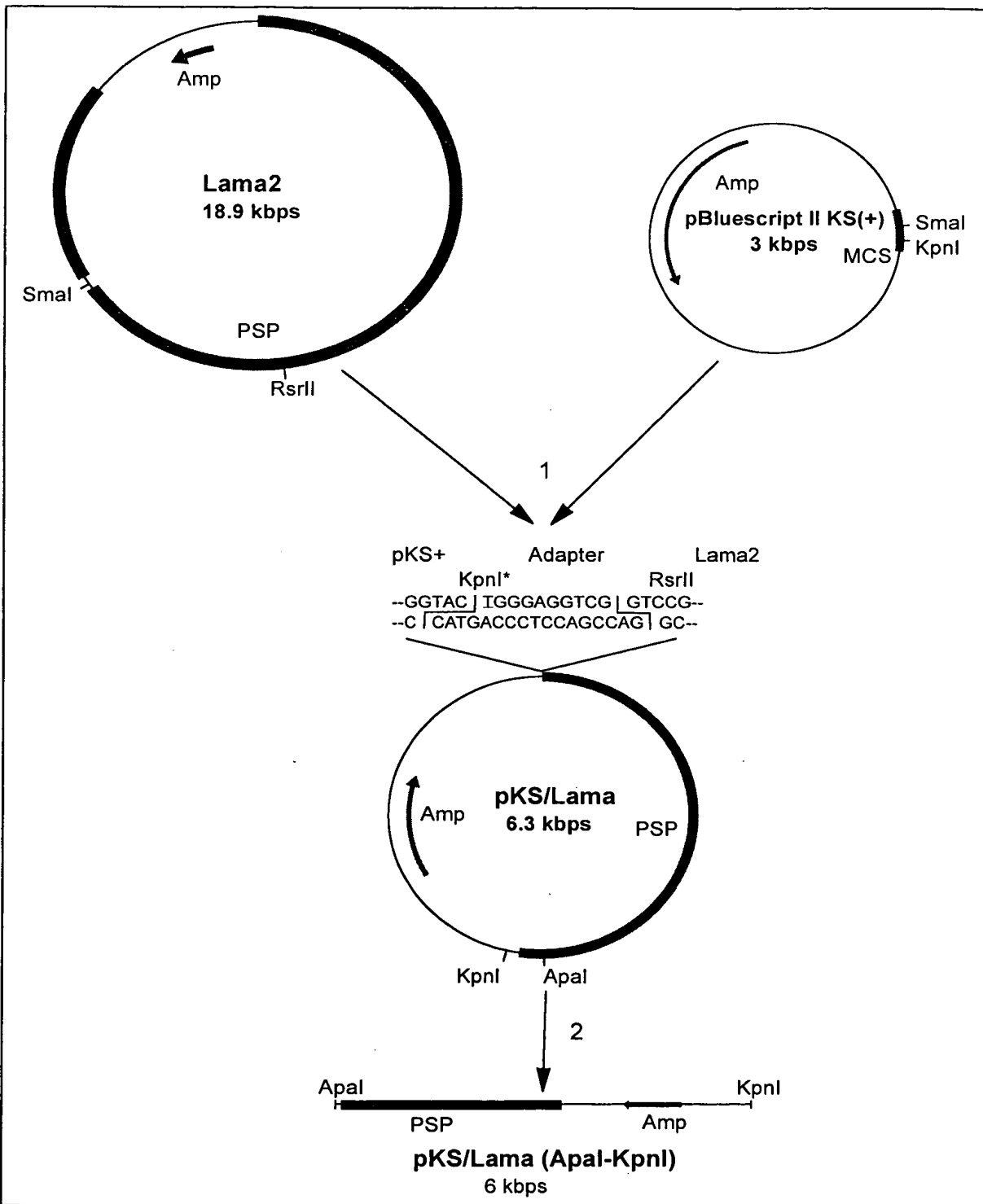


Figure 3

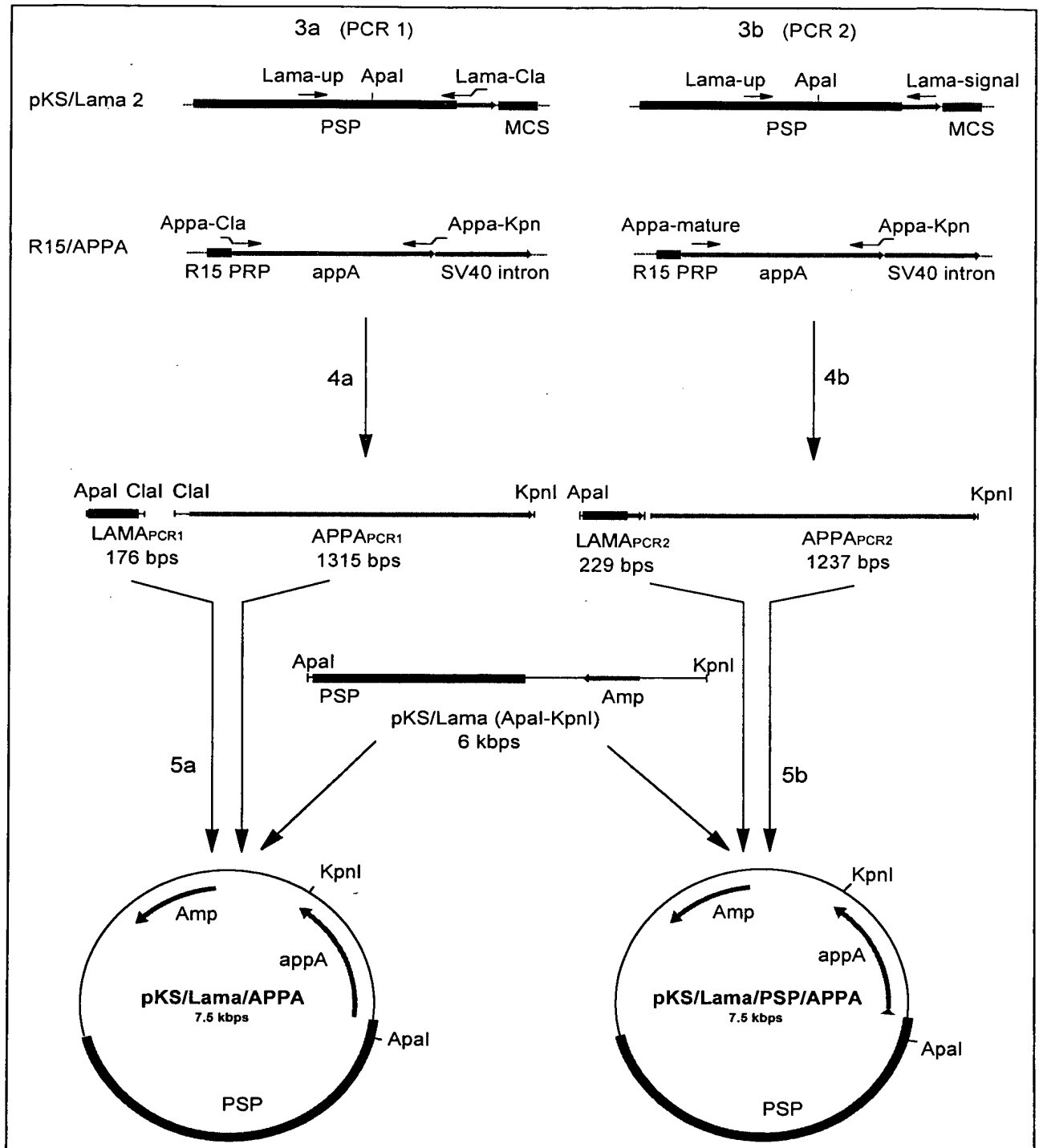


Figure 3A

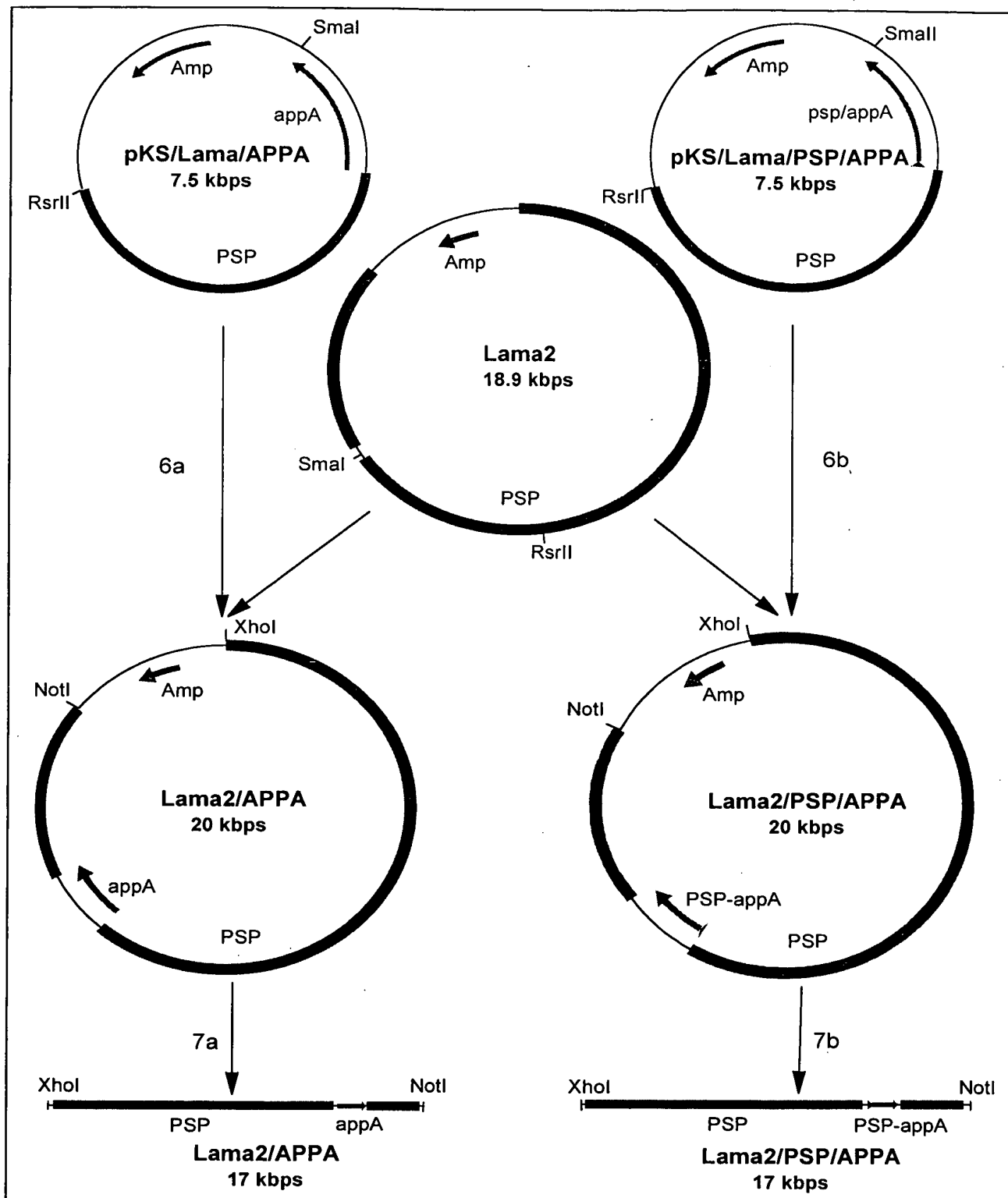


Figure 3B

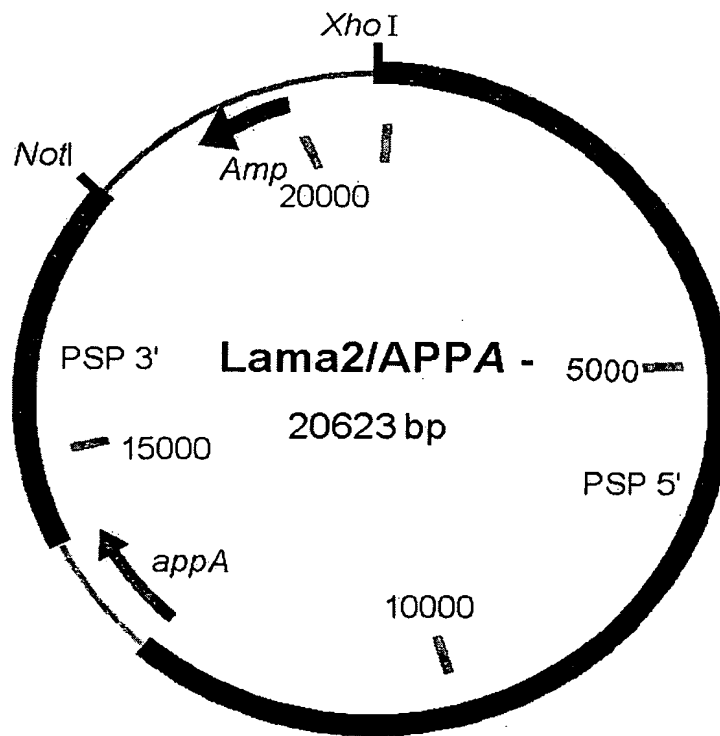


Figure 4. Schematic diagram of the Lama2/APP A construct.

Figure 5. The nucleic acid sequence of the Lama2/APPA plasmid (SEQ ID NO: 1)

```

LOCUS      Lama-appA  20623 bp    DNA    CIRCULAR SYN      17-JAN-2000
DEFINITION Lama 2/APPA transgenic construct
ACCESSION  Lama 2-appA,
KEYWORDS   parotid secretory protein; acid glucose-1-phosphatase; appA
           gene;
           periplasmic phosphoanhydride phosphohydrolase; artificial
           sequence;
           cloning vector
REFERENCE  1 (bases 1 to 20623)
AUTHORS    Golovan, S., Forsberg, C.W., Phillips, J.
JOURNAL     Unpublished.
FEATURES
DEFINITION M. musculus Psp gene for parotid secretory protein.
ACCESSION  X68699
VERSION    X68699.1  GI:53809
SOURCE     house mouse.
ORGANISM   Mus musculus
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
           Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 3777 to 5332;)
AUTHORS    Svendsen, P., Laursen, J., Krogh-Pedersen, H. and Hjorth, J.P.
TITLE      Novel salivary gland specific binding elements located in the PSP
           proximal enhancer core
JOURNAL     Nucleic Acids Res. 26 (11), 2761-2770 (1998)
MEDLINE     98256451
REFERENCE  2 (bases 7147 to 12653; 13952 to 17731)
AUTHORS     Mikkelsen, T.R.
TITLE       Direct Submission
JOURNAL     Submitted (07-OCT-1992) T.R. Mikkelsen, Department of Molecular
           Biology, University of Aarhus, CF Mollers Alle 130, 8000
           Aarhus, DENMARK
REFERENCE  3 (bases 7147 to 12653; 13952 to 17731)
AUTHORS     Laursen J, Hjorth JP
TITLE       A cassette for high-level expression in the mouse salivary glands.
JOURNAL     Gene 1997 Oct 1;198(1-2):367-72
MEDLINE     9370303

FEATURES             Location/Qualifiers
           source          1.to 12653; 13952 to 17731
           /organism="Mus musculus"
           /strain="C3H/As"
           /db_xref="taxon:10090"
           /chromosome="2"
           /map="Estimate: 69 cM from centromere"
           /clone="Lambda YP1, Lambda YP3, Lambda YP7"
           /clone_lib="Lambda-PHAGE (Lambda L47.1)"
           /germline
           /note="Allele: b"
misc_feature      3777-5332
           /gene="PSP"
           /function="salivary gland specific positive acting
           regulatory region"
enhancer          7147..8724
           /evidence=experimental
exon              11778..11824
           /gene="Psp"
           /note="exon a"
           /number=1
           /evidence=experimental
exon              12626.. 14190
           /gene="Psp"
           /note="exon b fused with exons h and i"
misc_feature      12644-12652

```

Figure 5A:

```

                /function=" consensus sequence for initiation in higher
                eukaryotes "
misc_feature 13952-13965
                /function=" M13mp18 polylinker"

DEFINITION  E. coli periplasmic phosphoanhydride phosphohydrolase (appA) gene,

ACCESSION  M58708 L03370 L03371 L03372 L03373 L03374 L03375
VERSION    M58708.1 GI:145283
SOURCE     Escherichia coli DNA.
ORGANISM   Escherichia coli
            Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
            Escherichia.

REFERENCE  1 (bases 12653..13951)
AUTHORS   Dassa,J., Marck,C. and Boquet,P.L.
TITLE     The complete nucleotide sequence of the Escherichia coli gene appA
            reveals significant homology between pH 2.5 acid phosphatase
            and glucose-1-phosphatase
JOURNAL   J. Bacteriol. 172 (9), 5497-5500 (1990)
MEDLINE   90368616

FEATURES              Location/Qualifiers
    Source              12653..13951
                        /organism="Escherichia coli"
                        /db_xref="taxon:562"
    sig_peptide          12653..12718
    /gene="appA"
    CDS12653             13951
                        /gene="appA"
                        /standard_name="acid phosphatase/phytase"
                        /transl_table=11
                        /product="periplasmic phosphoanhydride phosphohydrolase"
                        /protein_id="AAA72086.1"
                        /db_xref="GI:145285"

    mat_peptide          12719 13948
                        /gene="appA"
                        /product="periplasmic phosphoanhydride phosphohydrolase"

    mutation             replace(12659.. 12661,"gcg changed to gcc")
                        /gene="appA"
                        /standard_name="A3 mutant"
                        /note="created by site directed mutagenesis"
                        /citation=[3]
                        /phenotype="silent mutation"
    mutation             replace(13934..13936," ccg changed to ccc")
                        /gene="appA"
                        /standard_name=" P428 mutant"
                        /note="created by site directed mutagenesis"
                        /citation=[3]
                        /phenotype=" silent mutation "
    mutation             replace(13937..13939," gcg changed to gct")
                        /gene="appA"
                        /standard_name=" A429 mutant"
                        /note="created by site directed mutagenesis"
```

Figure 5B:

```
/citation={3}
/phenotype=" silent mutation "

DEFINITION      pBluescript II KS(+) vector DNA,
ACCESSION        X52327
VERSION          X52327.1  GI:58061
KEYWORDS         artificial sequence; cloning vector; expression vector; vector.
SOURCE           synthetic construct.
ORGANISM          synthetic construct
                 artificial sequence.
REFERENCE        1      (bases 17732 to 20623)
AUTHORS          Thomas,E.A.
TITLE            Direct Submission
JOURNAL          Submitted (20-FEB-1990) Thomas E.A., Stratagene Cloning
                 Systems, 11099 North Torney Pines Rd., La Jolla, CA 92037, USA
REFERENCE        2      (bases 17732 to 20623)
AUTHORS          Short,J.M., Fernandez,J.M., Sorge,J.A. and Huse,W.D.
TITLE            Lambda ZAP: a bacteriophage lambda expression vector with in
                 vivo excision properties
JOURNAL          Nucleic Acids Res. 16 (15), 7583-7600 (1988)
MEDLINE          88319944
REFERENCE        3      (bases 17732 to 20623)
AUTHORS          Alting-Mees,M.A. and Short,J.M.
TITLE            pBluescript II: gene mapping vectors
JOURNAL          Nucleic Acids Res. 17 (22), 9494 (1989)
MEDLINE          90067967
FEATURES         Location/Qualifiers
Source           17732 to 20623
                 /organism="synthetic construct"
                 /db_xref="taxon:32630"
CDS              complement (18967..19827)
                 /gene="Amp"
                 /product="b-lactamase"

BASE COUNT      5449 a   4847 c   4902 g   5424 t
ORIGIN
1  TCGAGAGTAT CTTTGTGTCAGC TGTGCCTCCA ACAAAGGGGT ACTGTTGCCC ACATAGAAAG
61 ATCTAAACTA ATTAATTAAT CCCTCACCCG CAAATCTTTC AGTCACTAAG TTAGCACGAT
121 TGTTGAACAA GTTCTCCAAA GGAGAGATAC AGATGAGTGC GTATAGGGTG GACCTGGCTG
181 CTGAGGAGAC ACCTGCATCT GACTAAGAAG AGCCACGGTG TTAGTTGAAT GGTGTGGAGT
241 AGGGTGTTTC TGTGGGACAG TAGAAAATCG AGAGGCATGT GCCGTTTAGT GAACTGATGG
301 AAGCTACCCC AAACGACAGA GATTGTCAGT CAGGCCAATC CGTTTCGAGT TTGATGGGCA
361 GCCGGACAGT GAGACAGACA CACCTACTCA GTTGGAGGAA GGATGAGAAC AATGGCCAGC
421 AGGGATTGAG AGACCTGAC AGGCGCAAGG CCTAACACA CACACCTACC ACCTCACTTG
481 ACAAAGCTGC CAAAGACCAA AGACTTGTTT TCCATTAGAA ATGACAGCTG GCTTGACCCG
541 ACAGCATAAT AAGCAGAGTG TACTCTGATT GGAGAACTTT AATGTGTTTC ATTCACTATT
601 ATAAAAGGAC AGTATTACAG ATTTTGTGTT ACACCTGCTG TACATGTGGG GCAGTGTGTC
661 TTTAAGTAGG GTAAAGTACT CTTTAAAAAT GGGTCCTAGA TATTTTTTCC TTTAACTCAA
721 GTCTCTTACT GTTTAAATGA TTTTATTTT GTTTAATATG GAGGAAAAAG AAGCGTAAAT
781 GGACAAATATA TATTTAGAGA AAGATGGTTA GCTGTCAGAA AAATATGCAA ATCAAAATCA
841 CACCAAGACT GCAGCACACC CCTGTGAGAT GGCTGTGATC AAGAAAATAA ATGACAATGA
901 GTGGTGGTGA AGATGTACTA AAGGGAAACA CACACACACA CACACACACA CACACACACA
961 CACACTGGAG CAACCACTGT GGAAATCAGT ATGAATGGTC CTCAAAAACC TGAAGATAGA
1021 GCGGGGCGTG GTGGCATACA CTTTTATTCC CAGCACTGGG GAGGCAGAGG CAGGTGGATC
1081 TCTGAGTTCC AGGCCAGCCT GGTCTATAGC ACAGGTTCTA GGACAGCCAG GGCTACACAG
1141 AAAAACCCTG CCTTGATTAA ACCAAACCAA ACCAAACCAA ACCAAACCAA ACCAAACCAA
1201 ACCAAACCAA ACCAAACCAAG ACCAAACCAA AACACTGAAG ATAGAACTTC AGTATTCCAT
1261 TCCTAGATAT ATACCCAATG GAGACTAAGT CAGCAAGACA CCTGCACAGC CATGTTCACT
1321 ACTACACTGT TCACCACAGC CAGGCTGTGG AACCAGCCTG AGTGTCATG ATAAATGAAT
1381 GGATAGGTAA CTTTCAAGGT AAATGGACTC TGCTGTGTAC ATGCCTCACA TTCTGTTTAT
1441 TCATTTTTCT TTATGAGGTG TCCATTCAAG AGTCACATGG TAGTTCATT TTCAGTCTTC
1501 TGAAGATACT ACACTGGTCC CCACAGTTTA CACTTTTATC AGCAGTGAAT AAGGGTTCCT
1561 CTATCCTTAC CATCATTTGT TGTAATTTTT CTTGATGACC CTCTTTCTGA CAGGGATAGG
1621 ATGTAATATC AGTGTGAGGA AGTACAACCT GTTTTCTAAG TATTTATTGG CCCCTTGCAT
1681 TTCTTCTTTT GAAAACTGTC GGTTCCCTGAC ATCTGCTCAG GTATTCATTG GATGTTGTTT
```

Figure 5C:

1741	CTTTGGTGTT	TGAGTTCTTA	TGAATTCTAG	ATGTTAAATC	CCTGCCTGTG	GTTCTCTCCC
1801	ATTCTGTAGG	CTGCCTCCTC	ACCCTGGCAA	TTGTTGTCCT	TGTTTTGCAG	AAACTTTTGA
1861	CTTCATGGAA	TCTCATTTGT	CAGTTTTCCC	TCCTCTGCTA	TAGCCTGAGC	TAATGCACTG
1921	GTTTTTACAG	AGCCCTGGTC	TATGCCTTTA	TCCTCCTCTG	GCAGCTTCGG	AGTTTCATTT
1981	CTTACATTTA	GATCTTTGAT	CCACTTTGAA	CAAGTTTTTG	AGCAGGGTGA	GAGATACGAA
2041	TCTAGTTCCA	TTCTTCCATA	TGTGATCCTA	GTTTACATAG	CATCGTTGGT	TGAAGAGGTT
2101	TTATTTTATT	TTTAAATAAT	GTGTCATAAA	AAACGAGGTG	GTTGTAGCAG	TGTGGATTTG
2161	TTTCTTTGTC	CTTTGATCTA	CAGGTCCTGT	TTTGTGTCAG	TCTCATGATG	TTTTATTGCT
2221	ATGGCTCTGT	CATACAGTCT	GAGGTCAGGT	ATTGTGATAT	ACCTTCAGTA	TTGCTCCCTC
2281	AGACTCAGT	TTGCTTTGGC	CAGGAGTCAT	CTTACTCAGT	GCTCTTAGAG	CTCCCCCAGC
2341	ATGTAGCTGC	TACTATTCTT	AGTTGATAAA	TCAGGAAACT	GGGGCTCAGA	GAGATTAAC
2401	GTCTTGAAC	ACTTCTGGGG	AGGTGAAACG	TGGAGACACT	AAACTGTGTT	TACCTGTAC
2461	TGCTCCAGTA	GCTGTCGGGT	GCTGGGCTAC	AGCAAAGCAC	CTATACTATA	TATTACTCAG
2521	GAGGTGGA	AACTCAGCCT	CCCTTGGGGT	TCCCAAGCTC	CCAGGTGTCC	AGTCACTGCT
2581	GGAAACCTCA	TGGAGTCTGA	AAGGAAGGTT	TGAGGTTACA	TGGGGCAGCG	ATGAGGAGCC
2641	TGGGGCTGGG	ATCTCCCAAA	CACCTGGATA	TCCAGATGCC	ACTGGGTCAG	GGGAGATTGG
2701	GAACAGAGTT	GGGATGTCCA	TGGACCTGTG	ACAAGGCCAG	GGCCAGGGGG	AGGATAACTC
2761	TGGCTTTACT	AATTGCGGAA	AGTCCTTAGC	TTAGCAGCAG	TTGTCTGGGA	GCACAGAGGG
2821	GCCTTCTGTA	AGAGGCTCAG	GCAGTGCCCG	TCTGTAGGCG	AAGGTCTTCT	CCATGTTCCC
2881	CATGGTGGTT	CTTGATGAAA	GAGACAGTCT	TTGGCTCCAA	ACTGGTTTAT	TGATTGTTCA
2941	TTGTGGAAAA	TGGGTGCACA	CCACCTTCTC	AGGGTGGACC	AGAGATCAAA	TACCTTTTGC
3001	AGGGAGGAAT	ATCTGGGAAG	GGACGCTTAC	TGGCTAAACC	CTCAGGGCCT	CTAGATACAT
3061	CATTAGCATG	GAGAACTCTG	TTCTGGGCTA	CATGACCACA	GGCCACATTT	CCACAAGCCA
3121	CATGTGGGAA	GTGTGGCACA	TGTTCTAGCG	CAGGAATCTG	GTAGGGAGCG	TGGAGCCACC
3181	TACCATCCCA	GGTGGGTGCC	TGGGTGGCAG	GGACCCTGAA	CCCGCTCAAC	CTTACCAAGT
3241	TTCCTGGCAG	GGTCCACTGT	CCTACACAGA	AGCTGGAGGA	GGTGTGAGGG	TTGTGTCTTT
3301	GTGGAATGTC	CCATGCTGCT	TGGGGCTCAG	TTTCTCCACC	TGTACCTCAT	TGGTTTGGGT
3361	ATAAAAAGTG	GGGATACTTT	ATTATTCTCT	GACTCGGTCC	TGAGGAAAAA	GCATCGTGCC
3421	AGTCCAGGAA	CCACACCCTG	AGGTTCCCTG	ACTGAAGGGA	CTCCCTAAGT	CTCTGGAGTC
3481	TCTCCCCTTC	ACAGAGCTGC	CAAAGTTAG	GTTCTTTTGA	GGATAACAGA	GCCATGCTTG
3541	GTAAGCAGAC	AACAGCATTT	GTTTACTCAA	CCTTCTTTTG	TCAGCTCCCT	CTTCATAAAC
3601	AAGTTGAGAC	ACCATGCTGG	CTTGAGGAAG	ACTTCTAAAG	CCAGACAAC	GTGCAAGGAA
3661	GAAGAAGAAG	GGGCAAGTGG	AGTTAGCCTG	GATGTAGCCC	TCAAAGTCTC	CAGAGACCAG
3721	CCATGAAGGC	TCAAGTGGAG	GGCAAGACCT	GCAGCAGCCA	AGCATCTGGC	AGGAGAGGAT
3781	CCTGGGAACC	CCTCTACCAT	GACACACATT	CTTCTGCAG	GTCACACTTA	ATTAGGCCATT
3841	TCTTATTTGG	ATCTATCATG	GTGTTCTGTG	CGAGATTAAT	GAGGTGTTAT	GCTGCGAACA
3901	GAAAGTTATA	TAAAAACAAG	TCCCCCCCC	TTGTCACTGC	TGCTAAGAAT	GTAGCAGAAA
3961	TTGTCTCAAG	TGTCTCTCTA	ATCAGAAACA	ATAAAGGTCT	CCTTGGATTCT	AAGCCCTCCA
4021	GTTTCTCTCT	TCCTTGCTGA	GCCTTGGA	CCCATAACAA	CCTCCTGGAT	GCTACAGCTC
4081	TGGGCAGAGA	CTCCAAGGTG	GGGAGAGACT	GATGGTACAA	AAGCAAAATA	CTTGTTTGGG
4141	GGTACACCCA	CTCCTCTGCC	TGTGTGGTTC	CTGCAGTCAG	TCCTGCAGAC	AGGCCCTCAG
4201	TGGGTCTTCC	ATGGGCAACA	CGCAGAGGGA	GGCAATGGAT	GGGAATACCC	ACACCCTGGT
4261	TAGTTTACCC	CGGCCATGCT	CTCTGCTCTT	CATCCCTCCT	CTGCCCTCTG	CCACGGCTTT
4321	CTCTGCAGGA	ATCATATCTT	CATATTGGCG	CACAGGTGTT	CTCCTCACCC	TAGCTATGAT
4381	GTTTACTTTA	GAGTGACCTT	AGCAGGCTG	GTGGGAATGA	GTTCTAGAAG	GCTCACGGAG
4441	ATGCTAGGGA	AGAAACGTCT	TCTAACTACT	GAGGTTACTA	AGTTCTTGGT	GGTTGTCTCT
4501	GCCTTTCCCT	TGTTAAAGTC	ACCTTGAAGT	TAGTGCAGAA	GAAATCAGAG	CCCAGTCACA
4561	AGGTAAATAT	GGTCTGAAG	ATTTCTTTTG	AGTGCCAGAA	ATCCATGACA	TTTCAAGAGC
4621	CCTCTTTGTA	CCTTAAGTCA	TTTGGGGTTG	TATCTTCTGC	TTGATGTATG	TGTGTGTGTT
4681	TATCAAAGAG	TGAGATGGTT	ACATAAGAGG	TGCTCTAAAG	GACAGAGAGG	ATTTGCAATT
4741	GTGGCATGTG	ACATCCTCAG	GCCTTGCTCT	GGTGCCAGGA	GGAAGTATG	CAGAAAAGAG
4801	TAAGAGGTCA	TTTCTTGGAG	GCTGTCACTA	TAGAGGAGAT	CTTACAGTGC	ATTCCTCCT
4861	CCAGGCCCTG	CCTGAGGATA	GACATGTGCT	GACTGCAACT	GAAACAGAGG	CTTGGGATGG
4921	AGAGTTAGGT	TCACAGAAGG	GAGGGTGGGA	GATGGATGCT	TGCTGGGTTT	TGGGTCTCAT
4981	CACAGCTCC	TGACCACCCG	GTCAGCCCAT	GTGCTTATTC	CATAGCTTTC	TTTTGCTATG
5041	TTTACTCAGT	GTGGTGTTTG	TTGGGACCCA	GCAGAAGCCA	GTCCCAGGCT	GACAGCTGTG
5101	GATACACAGG	GCAGCATGAG	GGTCTCTCAG	CTGAAGCAGT	CAGGCTGGCA	GAAGAGAAAG
5161	ACCAGCACAC	ATTCCTTCAA	CCAACTATGT	CTTGAAAAAC	AAACATATTA	TATCACATAT
5221	ATTGCATTTA	TGAGACAGCT	AAAAATGACT	CGGGTAGCAT	GACTCCAGGT	GGGGATATCT
5281	GCAAGTGCCA	TGAGTGGCAG	AGGGACAGCC	AATGTGAGGC	AAGAAGGAAT	TCTGGCTCAA
5341	CACAGCTTAG	CTCCCTGGTG	TTGGTTCAAA	CTTTGAGAGT	TTGACCACAA	GCACCTTTAT
5401	TTTGACATAT	TTAAACAGAG	CACAACTTTG	GGAAAAAGTT	TTCTTATGAA	AATTATCACA
5461	ATAAAGCTTA	AGGCATGACT	ACATTAAAA	GCCTTTGCAA	AGTATATGTG	CCCTCTTCCA
5521	CAAGATGGT	TCTATTGACT	GAGAAATAAT	GTTCAGGATA	AAGATCCAGG	AAGAAAAGAT
5581	CAGGGATAAG	TAAAATACTA	AACTCTTTTG	CAAAGTACAT	AGACCCTCTT	TCATAACAA

Figure 5D:

5641	GGGTTCTATT	GA CTGACAAG	CA CTGCTCAG	GAGTTGGGAA	AGAGTCTAGC	ATAAGCACGA
5701	TAGCCTGGAG	ACTCTAGTGA	GGTCTAGTCT	TACAGACAGC	AAAAATCACC	AGGTTACAAA
5761	CTACATTCAT	TTCCAGTTTT	CTGATCAGGC	ACAGGTATGA	ATCCCTTCTG	TTGAAGAGAA
5821	AAGTCCATGT	GTTTAAAAATA	TCTGGTTTTCT	CCAGTGCTAT	TAGCGAGAAG	ACTTGAGCCC
5881	TATACAACCTC	CCACCTGGAG	TGACATCCTG	TCTTCATGGT	ATATTACATA	CCTAGACACG
5941	CTCATCTCAC	AGACTTAGGA	CTTTGTCTTC	TGATCTCCAT	TTCTGATCCC	ACTTCCACCT
6001	TTGCCTTGAT	AGTGTCA TTT	TCTTCACTGC	CTTGGTGACA	ACCATGTTAT	CCTCTGTGTA
6061	TTTGAGTGTT	ACCATT TTTCA	GATTTTACCT	GTATGCAAGA	TCACACAGTC	TTTGTCTTTT
6121	TGTCTGGATG	CATGCTAATC	TCTACACAAC	AACCCTTCCC	CGTCACTCAG	ATCTTCTCTC
6181	ATTAACACAT	ACATGGTGCT	GAAGAGGCTA	GGGAGCTTCC	CTTCAGTGGG	GAGCTAGCTG
6241	GCTATTGGGC	CTTTTTGACT	GTCCAGGAAG	GCCCCCAATT	GCTGAGACAA	GAACTTAGAT
6301	TCTTCATTAT	TGACTCTAAC	TCATGTATCA	AGCAGAAGCT	AATGAATAGT	TATCAACAGG
6361	ATCAGAGGTT	CCAGTGTAAG	ACACTTTGAC	ATGAAAGAAC	GGAGGAAGGA	CAGATGGATG
6421	CATAAAAGCA	GGACCACTGC	CCCAGGAAGG	TCCTGGAAAC	TGATGCAGGG	CAAAGGACAG
6481	GTTATAAAACC	AAATCTTAGG	GAGTCAGGAA	GAGCACAGAG	GAGCTCAACC	AACTGACCAC
6541	TGCTTAGGGG	CTACCAACCC	AATCCTCCCT	GTGGGAACAG	CTAAGCTATC	AGCCAAGGGT
6601	AATAACAGAG	CAGGACCTGT	GGATGACATG	GAGAGCATAG	GGACCCTGGG	TCCAGCCTTT
6661	AGCACCTGCA	CTCTCAGGAT	ACTCCACCAT	TGTGTCTTAG	AGAGCCTAGG	GATACTGGGT
6721	CCAGCCTTTG	GTACCTTCAC	TCTCAGGGTA	CCCCATCACT	GTGTCTTGGA	GAGCCTAGGC
6781	ACCCTGGGTC	CAGCCTTCAG	TACCTGCGTG	CTCAGGACAC	CCCACCATTG	TCTGTGCCCC
6841	CGTCTCTTCT	TCCTCTTCCT	CCCTTTTCATT	GTCTCTTCTC	TGTTTTCTTT	TTGACTCTCC
6901	TTCCCTCTCA	CACCTCACT	CTAGTTCTCC	CCTTCCCTCT	CTGCATCACC	CTATTCTCTC
6961	TGTGGTCCCT	CCACTTTCCCT	TTATCTCTCA	TGCTTCTCTC	CTCCCTCAAA	TACTTGTTCAC
7021	CCACTATACT	TCAGGGGCCA	GCTCTAGTGA	CAAAGCTGTT	AATAGCAAGA	CTCTCAGATC
7081	TCCAACGGCT	CAGAGGAGCC	AGACCCACCA	AGAACTCTCT	CCAGGTCCAA	TTTCAGGTTT
7141	CTTCGAAAGC	TTTCAGCAAA	TGCTCAGGGA	ACATGCCACT	AACAAGAAGA	TGCAAAATTC
7201	AGTTGAGAGT	GGGAAAGGCC	CTTGCGTAGG	TCCCATCTTC	CAGGCCAAGG	TCAGAGGGGC
7261	TCTGTGTAAT	CCGGATTGAC	AGGGCTCAGA	ACAATGTTTT	GTTTTTAAAG	TTTATTTATT
7321	TTAGGTGTTA	GTGTCTTTGC	TTGCATGACC	TTATGTGCAT	CATGTGTGTG	CAGGTTCCCTG
7381	ATGACAGTAG	AGGAGGGCTT	TGAATCCCTG	GGGATAGGAA	GTTACAGGAA	ATTATAAGCT
7441	GCTTTGTGGG	TCTTCTAGCT	TTCCCAACAG	AAGTGAATGC	TCTTCACCAC	TGAGCCATCT
7501	CTCTAGGCCC	AAGAGACATT	GCTTTATGGA	TATAATTGTG	TGTGTGTGTC	AACATTGAGG
7561	AAAGGGAAAT	AAAAAAAAAA	CTTCAGCCGC	TAAGGTGTGA	CAGTTTCACT	AATTGCTACT
7621	TTAGTTGTG	ATAAAATGGC	AGGTGCTTCA	ACATTTATAT	ATACAAAAAC	TTCCCTGCTG
7681	GTGGTTCAAC	TGTGAGAACT	GGGGTAAGTG	GGTGAGTTCT	CTTTTTCTGT	CTCTGTCTCT
7741	GTCTCTCTCC	TTCCATTCTT	TCTTAAAGGA	AATAAACATT	GCAGCTGGGT	TATAGCTCAT
7801	CAATATGGAA	GTTACAGAAG	TGAAAAAAGG	CATTGCCTTG	GTGGGTGGTG	TTACCAGCTG
7861	ATTTTGGGTT	GTCCTGCAAG	GAGGTCTGGG	GACTGGCTGC	TCTGTCTCTG	TCTGTATGAG
7921	TGAGGGAAGT	CTGGGGAGCA	GATTCCTTAA	CCTTCAGCCT	GGCCTGGTTC	CTGAGTGAAC
7981	CCAGCCTCTC	TGGTCCTAGT	AGCTTTTTC	AAACAGGAAT	CTGAGTGGTG	ACAGGGAACA
8041	AGTACCAGCC	CATTGCTTAA	GTGCCAGGGT	TAGTGAGGGC	AGGAAGCTGC	CATAGCTGGG
8101	ATTAGTAGTT	GTATTGGATG	TAGGAAGTCC	TATCCTGGGA	CAGCTAATCC	TTAATGCTTC
8161	ACTGGAGATT	TTCAATGAGA	AATTTATCCC	ACGGCCCATC	TGGCCCCATC	CTTTTGTCTC
8221	CAACAGCTAA	GTATTTTCCA	TTAGAGGAGA	CTTCTGTGTA	ACTTGATGGA	TGCTCATTCC
8281	AAGGTGACTT	GGGGCAGTCA	GTACAGACTT	GGGATGACCT	CTGACAGCCT	AACCTCTCCC
8341	CAACAAGGGC	CCTCTATGTT	TGCTATGTAA	TGTAATGTCA	GACATTGTCA	GGAGTGTCCG
8401	CAGCACAGCC	TGCCCAGTGT	GAGGGCTCTC	ATAGGTTTCC	CACTGTCTTA	TCTACACAGG
8461	GATAACGAGG	AGGTAAGCTG	CAGTTCCCAG	TCTCACTTCA	CAGAGGAAGA	GATAACCCCA
8521	TCCCAGGTCA	TGTAGCCAGC	AGTGGAAGA	ATGAGGATT	GAACCTCAGG	CTTCCAAGTC
8581	CCATTGATAG	CATCTCCTCA	CAAGTCCCTT	GCCACCCTCA	CGATGCCTTA	GACACTTGCC
8641	TGCCCTTTAT	ACTAAGGAGA	TGCAGGTACA	AGGGGTTTAC	CCATGTAGCA	GCTGAGGCAG
8701	CTGGGGATAG	ATACCAGCAG	CAGGCCTGAT	GTCACCACTC	TAACCTCCAGC	ATCCCCAGTC
8761	TGTGTTCCTG	GAGTGTGAAA	ATCCCTACTT	AACAAGATTG	TGCAACAGTC	CTTGGCTCTG
8821	TGACCCATAG	CTGGAACACG	GATTCTCATT	GATTTGTGGA	ACATGGTGGC	AGCCAGCCAA
8881	AAAGAGGGTC	TGCATACAGA	AGACACGTGT	GGCAAGGCCA	CAGCAGACTC	TGACTACCTT
8941	AGCTTACAGA	ATTACAAGGT	CATAATGTCC	TCTGCTTTGG	TCACCTCATG	TTAAGGACAG
9001	GCCCTAATGA	AGATGGGGCA	GAAGACTGAA	GGAATGGCCA	ACCAATAACT	GGCCCAACTT
9061	GAGACCCATC	CTACAGGCAA	GCATCAATTC	CTGACACTAC	TAATGATACT	CTGTTATGCT
9121	TGCAGACAGA	AGCCTAGCAT	AACTATCCTC	CGAGAGGTCC	ACCCAGCAAC	CTGATGAAAC
9181	AGAAAAAGAT	ATCCACAGGC	AAACAGTGGA	TGGAGGTCAG	GGACTATTAT	GGGAGAGCTG
9241	TGGGAAGGAT	TAAAAACCCT	GAAGGGGATA	GGAACCCAC	AGGAAGACCA	ACAGAGTCAA
9301	CTAAGAGACC	TGTGGGAGCT	CTCAGAGACT	GAGCCACCAA	CCAAAGAGCA	TACACAGGCC
9361	GGTCCGAGGC	ACCTGGCACG	TGTGAAGCAG	ACATGCAGCT	CAGTCTCCAT	GTAGGTCCTC
9421	CAATAAGCGG	TAGCCTGACT	GCAGTATCCA	ATCCCCAACA	GGGCTGCATA	GTCTGGCCTC
9481	AGTGGGGGAG	GATGCCCCTA	ATCCTGCAGA	GACTTGATGA	GTGGAGAGCT	ATCCAGGGGG

Figure 5E:

9541	AACCCACCT	CTCTGAGAAG	GGAATGGGGA	TGGGGGAGGG	ACTCTGTGAA	GAGGGGACAA
9601	GGACAAACAA	GAACCTCAAA	TAGGTCAGGC	CCTAAAGGCT	TGCTAAGTAG	CAGTGGCCCA
9661	GCTCTGTCT	GTTCTCAGC	CCAAGGCTCA	GCTCCACCT	GTTTCTGTGT	TTTTCTGGCT
9721	TTTCATGGGC	CTAGGACTTG	GTGACCAGTT	CAAACAATGG	GGCCTGTGGA	AGACACAATA
9781	TACAAGACTA	GGGACATTCC	TGTTCTGCTG	ACTATCCATA	GCCTGATGTA	GGTGAAGGA
9841	CCCAATCACT	GGATTCTTAC	CCTTGACACAA	CCTTGACAGC	TGAGGGCCTC	TCAGAAACCT
9901	ATTTCTTCCA	CTGAAAAATG	AGACTCTCAA	ATGAACGTCG	TGACAATCAT	CAGGCTTATT
9961	AAAGAGGTGT	ATCTAACCTG	AATGGCAAGC	AGACAGCAGG	CAAATGTCTG	TATCAACCTC
10021	TAGGAAGGAC	AAGAACTGCT	CACTGCTGCC	CCCCAGGAGG	CCATTTGCTG	AAACAGCTGC
10081	TCTCCTGCTG	GTGCACAGGC	CCTGCCTTCT	CATTGCAGCC	ACAGCCCCTT	CCTGTCTGAA
10141	CCTCCTGTCA	GGTCACTGGG	AAACAGATCA	AGATGGAACA	GGACAGCTCC	TGATGGTAAA
10201	TAAAAAACAG	TGGTCATGGC	TATTCATAGG	GGTTTATGCT	TCTTCAGTCC	ACACTGTGAA
10261	GAGCTGTGGG	CATGAACCAC	AGTGTTTCAG	GTAGAGTTGG	GGTTCTGAAA	TTACAGTGG
10321	GGTGAGCTCA	GTAAATGTGA	GCTGGAGGTC	ACTCGTGAGA	CACACAGTCC	TGCTGCTTCT
10381	GTTCCCAATA	TCCTGAGGAG	ACGACACATC	TACTTTGTTC	AGAGGCCACA	GTCTAGTTGA
10441	CCTGAGAGTT	ACCAGTTTCT	TATTTGTGTG	TGTGTGTGTG	TGTGTGTGTG	TGTGTGTGTG
10501	TGTTGTTTCT	GTGTGAGTGC	AGGTGCACAT	ATGATAGCGT	ACACGTTGAG	GTCAGAGGAT
10561	AACTATCAGG	CGTTGTCCCC	TCCTACTTTT	CCTCGGACTC	TGGAGAACAA	ACATGGGTCC
10621	TTATTCCAGG	GGAGCAAGTC	GCTGTTGGCT	GACACATCTT	GCTCACATAC	ATTTTACCTA
10681	GACAAATGAG	CCTCCATCAG	AGTATTACTT	TAGCTCCTCA	CCGATGGCAA	TGCACCACCT
10741	CTCTACCCAC	ATAGGAGTTG	GGTCTCCACA	CACCCCCACA	CCCCCTTCAC	CAAAACGTTT
10801	TCAGTTACTT	TATCTGGTAA	AGTTCATCAG	AGAATGAAGC	CAGTATTAAG	AACATGGAAT
10861	CATTTGGGAA	CCTGGATCTA	GCAATACCCC	ACCCTAGATG	GAGTTGCTGA	GTTTTCACCT
10921	CAGATTATAA	TTCCCCCTTA	GCTTCTATGG	TTTATTCTGA	AACCAGGGGA	ACTCGATTCC
10981	TCCCTTTGGA	CCACAGACAT	CCTGGCTTGT	GAATTCACAT	GTCATCTACT	GTCATCTACT
11041	TGGTAGTATG	TGGCTCACAG	AGACACACTA	CAGTCATGGC	CAATGTCAAG	GTAGGACAGA
11101	TGTGAATCAT	TCCCCAGTC	CTGCTGTTTT	CATGACTAAC	CCTCCTCAGC	ACAGTGACCA
11161	TGAACCTACT	TTTCCCTCC	TTTTATTTTT	AGAATTGCTG	GAATTTTCTA	TTTTGAGAAA
11221	TAATAGCCTT	GGGCAGCATT	AAACAAAATC	ATCTAGAAAG	CTGGTTTTAA	ATACAGATGG
11281	TTGAGTCAGT	GAAAGAGTGA	GGAATGTCTA	TATTGGCCCC	TCACAGAGGC	TGGCTCACTC
11341	CAGCAGAGGT	GGTTGAAGCT	CTTGGACACG	GGTCAGGTGC	ATAGGAAAGG	TNGTCTGGGA
11401	CACTGAGAAC	CACAATTGAA	CAAACAGAAC	TGTTGGCTTT	TTTTTTTTTA	AATGAGTTCT
11461	CAAAAAATGA	CTGGCTAGCT	TAGGCAATA	CTTCGAGCCA	ACCCAACAGA	ACATCTTCC
11521	ATTGATTCTT	TCTGGATCTT	CTTTCTAGAC	AATACTGAAC	TGACCCCTTG	TTGGCAGTCT
11581	CAAGTTTGAC	AACATAGGGC	TTTGAACCTG	GCACAAGGTC	CATCACTGTC	ACCCAAGCAT
11641	CCTGGGTGAC	CTTTGGGTG	GAATATCTTG	GCTAACCTTA	GATATTTTCT	TTGGAGTATC
11701	TTTAGAACAT	CCAGGAAATA	GGGCTTGATT	CTCATCCTGG	GACCACAATA	TAAGTCACCC
11761	TAGAATCCCA	GGAGATCGTG	CAGAGAAACA	AGGATCTCTC	TCGTGTGCAT	CCTTCTTCAA
11821	AGCAGTGAGT	AGTGACTCCA	CTAAACTGAG	TTCCCATCTG	AGAGTCCACA	GGAGGCTTTG
11881	GGCAAGAAAG	CAGAGGGAAG	GCACGTGTTG	TGTTGGTAAA	GTTTTGACTC	TAACAAATTT
11941	GAAGACATAG	ATGACATTGT	GTCAGACTAA	CAACAACCTA	GACTCATGTG	GGTTCTGTTT
12001	AGGGATCAGA	TTTTATTCAT	CAATGACTTG	TCTTAGTGTA	TAGAGAAAGG	CTTCTACTG
12061	GAGTGATAGG	TCAATAATGA	CAGAAGAGAT	AGCTATTTCC	CCTAGGGACT	GTGCTGCTCC
12121	AAGTTTGGTG	GAGAAAGGCA	GTGGGGAACC	TAGATGTGCT	CTCTGGGGAG	GGGGTCTGAA
12181	CTGGGCTTCA	TAGAAGGTGT	GAAGTTTTCG	TGAAACATCT	AAACAGAATT	ATAGCTTAGG
12241	AAAGTGAGCA	GGCAAGGCAG	GGAATGTGTT	GCATATGTAT	ATGTACATGA	ATATATTATG
12301	TTATAGATAC	ACACACATTT	GAACCTCATT	TGCAGATGAC	AGAAAATAGG	TTATTTTGCC
12361	TCTCTTAACT	GCTAAGCACA	ATGACTTCCA	GTTCCATCCA	TTTCTGAAA	TGCCACAATT
12421	TCATTTTTC	TTGTGGCTGA	ATAAAATTCC	ATTGCAGACT	GGGCCCTACT	TCATCCTACT
12481	CTGAGGGCAG	GCATATCCCC	TGGTCCCAT	TCTTACCTAT	TGTGAAGAGA	AGTGCACTG
12541	TCTTGTTGAA	AGGCAAGCGT	GAGAGAGGCA	GGCACTAATT	GTGGGTTTTT	GTTTCTTCTT
12601	CCTGCTATGA	CTCTCCATTT	GTCAGAACCA	AAGATCGATA	AAAGCCGCCA	CCATGAAAGC
12661	CATCTTAATC	CCATTTTAT	CTCTTCTGAT	TCCGTTAACC	CCGCAATCTG	CATTGCTCA
12721	GAGTGAGCCG	GAGCTGAAGC	TGGAAAGTGT	GGTGATTGTC	AGTCGTATG	GTGTGCGTGC
12781	TCCAACCAAG	GCCACGCAAC	TGATGACGGA	TGTCACCCCA	GACGCATGGC	CAACTGGCC
12841	GGTAAACTG	GGTTGGCTGA	CACCGCGCGG	TGGTGAGCTA	ATCGCCTATC	TCGGACATTA
12901	CCAACGCCAG	CGTCTGGTAG	CCGACGGATT	GCTGGCGAAA	AAGGGCTGCC	CGCAGTCTGG
12961	TCAGGTGCGG	ATTATTGCTG	ATGTCGACGA	CGGTACCCGT	AAAACAGGCG	AAGCCTTCGC
13021	CGCCGGGCTG	GCACCTGACT	GTGCAATAAC	CGTACATACC	CAGGCAGATA	CGTCCAGTCC
13081	CGATCCGTTA	TTTAATCCTC	TAAAAACTGG	CGTTTGCCAA	CTGGATAACG	CGAACGTGAC
13141	TGACGCGATC	CTCAGCAGGG	CAGGAGGGTC	AATTGCTGAC	TTTACCGGGC	ATCGGCAAAC
13201	GGCGTTTCGC	GAACGTGAAC	GGGTGCTTAA	TTTTCCGCAA	TCAAACCTGT	GCCTTAAACG
13261	TGAGAAACAG	GACGAAAGCT	GTTTCATTAA	GCAGGCATTA	CCATCGGAAC	TCAAGGTGAG
13321	CGCCGACAAT	GTCTCATTAA	CCGGTGCGGT	AAGCCTCGCA	TCAATGCTGA	CGGAGATATT
13381	TCTCCTGCAA	CAAGCACAGG	GAATGCCGGA	GCCGGGTG	GGAAGGATCA	CCGATTCACA

Figure 5F:

13441	CCAGTGAAC	ACCTTGCTAA	GTTTGCATAA	CGCGCAATTT	TATTTGCTAC	AACGCACGCC
13501	AGAGGTTGCC	CGCAGCCGCG	CCACCCCGTT	ATTAGATTTG	ATCAAGACAG	CGTTGACGCC
13561	CCATCCACCG	CAAAAACAGG	CGTATGGTGT	GACATTACCC	ACTTCAGTGC	TGTTTATCGC
13621	CGGACACGAT	ACTAATCTGG	CAAATCTCGG	CGGCGCACTG	GAGCTCAACT	GGACGCTTCC
13681	CGGTCAGCCC	CATAACACGC	CGCCAGGTGG	TGAACTGGTG	TTTGAACGCT	GGCGTCGGCT
13741	AAGCGATAAC	AGCCAGTGGA	TTCAGGTTTC	GCTGGTCTTC	CAGACTTTAC	AGCAGATGCG
13801	TGATAAAACG	CCGCTGTCAT	TAAATACGCC	GCCCAGAGAG	GTGAAACTGA	CCCTGGCAGG
13861	ATGTGAAGAG	CGAAATGCGC	AGGGCATGTG	TTCGTTGGCA	GGTTTTACGC	AAATCGTGAA
13921	TGAAGCACGC	ATACCCGCTT	GCAGTTTGTA	AGGTACCCGG	GGATCACAAAC	TTGCCCTCTG
13981	AAGAGGAAGA	ACAGAAGGAT	GCCACAACCT	TCCTGCTGGC	TACTCTCCAG	TGGTTTCATC
14041	TTACTTCTGA	TGGCATTTC	CTCTAGAAAG	TGCTACTATC	ATCCACACAT	TTCTACCTGA
14101	GACCACCCAA	AGGACCCTCC	CAAATTCCT	TCCTCTCTGA	GTAAGTCTCCA	CACCTGTTAC
14161	CACCATCCCA	GAATTAAAA	CCTAATGCGA	CTCTGGCGTG	TGACTTGCCCT	CAGTCCCTTG
14221	AATAAGAGTT	GTTGGCAGTG	CCAGGCTGGG	TGGCGCACGC	CTTTAATTCC	AGCACTTGGG
14281	AGGCAGAGGC	AGGCGGATTT	CTGAGTTCGA	GGCCAGCCTG	GTCTACAGAG	TGAGTTCAG
14341	GACAGCCAGG	GCTATACAGA	GAAACCCTGT	GTGAAAAAAC	CAAAAAAAG	AAAAAAGTT
14401	GTTGGCAGAG	TGTGGGTAT	ATACCAGGTG	GAGATTTCAA	ATGAGTGGCT	GAAGCTGTAG
14461	CCAGAAGGAA	CTTAGAGGAT	AGCTCATAAC	TTAAAAAGAA	ATGTAGAGAG	TAGCAGAAAC
14521	ATTGAGAGAG	TGGGCACACA	GCCACTGTGT	GAATGTGGCA	GAACACAATC	CAGCCAGCTA
14581	TACATGCATA	AGTGTATATT	GGCGCCATCC	TGACTGATGA	GACACAGGAA	AACAGATAGA
14641	CGGGGTTAGG	TGGCCATGGC	CTTTCCTGCC	TGCCTCTTCC	TAAGGGTCAT	CTCAAGACCT
14701	TATGCTCTCT	TAACTCTTCC	ATTGCTACTT	AGCTTCTAGA	TATCACCTCC	AGATTAGTCT
14761	CCTTGGGTAC	ATCAGTGATC	CTGGTGATAT	CCAGGGCTTC	CTGATTCCAT	CTTTGTCATA
14821	GAGGCTGCAA	CTAAGAGGT	CTTCTTAATA	CTTCACACCC	TGATGCCAAA	AGGAAGACAC
14881	AGAAGTTCAC	AGAGGTGAAG	TGATTCTATG	AGGACATACA	GTGAGCAAGC	ATCAGGGTCC
14941	GGATTATCTG	ACTCTACTCT	AACTTTTATG	TAAATGTGCT	TTATGCCATT	AACACTGTCA
15001	TTCCTGTGCT	TCAGCTCTGG	GAGACTCCCA	AGCACTCTTA	GGCACAAGCC	ACAATTAAGG
15061	GACTCTGACA	CTCTGCATTG	ATTAATTAGC	ATGGTGGTCT	CTATGTTTCC	AGATTTCATG
15121	TGTTTTCACT	TTCCATATAG	GCTATGAAGG	GTGTGAGGAA	ATTTTTTGGG	GACAGAATTG
15181	GAGGCAATCC	ACCTCTCTCA	GGAAGCCTCT	ATCTGGAAAA	GCTTACAAC	CAGGAGCAGT
15241	AACTGTAGGC	CCAGTCCTTG	GTGTCCAAAA	TGGGTTTTAT	GGTTTGAATC	TGCAAAGCCT
15301	TCCATGTGCT	CAAAGGTTTG	AACATGGAGC	CTCCTCCTGG	TAACACTGTA	TGGAGGCTT
15361	TTGAGACTGG	ATGCTCTTTG	GTCCCATGTT	TTGCTACATC	ATCTGTCAAG	ATATGACCCA
15421	GGCATGCTAC	AGCTACCAC	AGACTATGCC	TCTCCAGCTT	TCATGTTCTC	CCACCATGA
15481	TAGACTTGTA	TCTCCTAAAA	ATGGAATCAA	AGCAAACCTT	TCCTGCATTA	AGTTTTTTTT
15541	TTTCTGTAA	GTGTTTGGTC	ACAGGGACAA	GAAAACACTC	AATACAGATA	ATTAGTACCA
15601	GAGTTGAGGT	TCATTGCTCT	AGCAAGTTGG	ATCAAATTTT	TAGGGCTTTG	GAAGTGAATT
15661	ATAAGAGACA	TGTAGAAGAG	TCTGAAGCTG	TGGGCTACAG	AAGTGTCAAC	AGTTTTTAAG
15721	AATAGTTTTA	TACACCATGG	GAATTGTGAA	AATCAGAAATG	CTCACACAAA	GGCAGCAGG
15781	AAAACGTGAG	CATGTGGCGT	GTGAGAGGGC	ATAAGAAGGA	ACCTAGGGGG	AAATGAGCTA
15841	GAAGCCATTC	GGCTACGTTA	GGGAACGTGT	GTGGCTGTGC	TTGGCCCATG	CCCTGGCAAT
15901	CTGAATGAGG	CCAAATTTTA	AAGGAGTGGA	CTAACTCGAT	TGTCAGAGAA	AATATCAAGA
15961	CAGACCCACA	CTCAGGCTAT	GCCGTGTTTG	TGACCGACCA	GCTACTCTTA	GCCAGTCTA
16021	TTGTGAATTT	CCAGAGCAAT	TATCAGAGTA	TGAAGATACA	TACAGTTTAT	TGAAGTAAAG
16081	GGTGTGGGTC	CCTAAGTGGA	TGGTGCATAA	ATCTATGTAG	GTGATGCCTA	AGTGACACTT
16141	GATAATCCAA	AATATCAGCA	ATGTGGAATG	TCTTCCAAGG	AGACCTGTAG	ACACACATTT
16201	TGAACCTTTG	CTCATGGCTG	TAATAAATAG	CTAGCTAGAA	ATCATTTCTT	GAAGAGGTTA
16261	GTCTGAGTTA	CGGTTCCAGG	GCAAACATTC	AGTGATGGCA	AGGAAGGCAT	TGCAGTCAGG
16321	AGCCAAAGGT	CAGCTGGTCA	CATTGCATCA	AGAGTAGAGA	GTCAGAGTGT	GAGTAGAAAG
16381	AGGATACAGG	TTATAAAACC	TCACGTGCCA	CTCTCAGCAA	TCCATTTTCT	CCTAAAAGGC
16441	TTTACCTTCT	AAAGATTTTA	GTCTTCAAAA	CCAGTACCAG	TAGCCTGGGA	ACAAAAGTTG
16501	AAACAAATGA	GCCTTTGTGG	GGCATTTCAC	ACTTAAAAAC	GGGCATCACC	TAGGAGGAGC
16561	CCTGTGTGCA	GTAGGAAGTG	TGGCCTCTGT	GTCAGGAATG	CTCAGGCTAA	TAAGGGGTCC
16621	TCTATCTGAG	GGACCCATATG	AAGATTCAAC	AAGTAGTTGT	GAGAATTCCC	TGTAAATGGA
16681	TGCTACCAAT	TTGACATTTG	TAGACCTGCT	ATTGTGTGCT	TCTTTATTGG	GCTCTCCCAT
16741	CTCCCAACTT	TCCAACCCAT	ATTCCACATT	AATCCCTTCC	ACCACCATGC	AACACTAGGT
16801	AGGAGAGAAG	GAAGGTTAGA	AGAGAAAGTG	GGTATAGATC	TATTTAGACT	ACTTCCTGCT
16861	GATTAGGGGC	AAGTCCAATC	GTCATTGTCA	GGATACCTCC	AACCAGCAAC	CAGCAAAACA
16921	GCAAATCAGA	AACAGCAAAA	GCAGCCAAAC	AGGCAGCACT	AACCAGCAGG	ATTGGGGTCG
16981	GTAGCGTGGG	AGCAGTCACT	ACTGGTCTTC	TCATGGCTTT	GGCATTAATA	CTCTCTCAAG
17041	AAATTCCGTA	ATTTTTTCCC	CACCACCTGA	AATTCCGTAA	TTTTAAATGC	AAACTATCTA
17101	CAGCTGGCAA	AAATCACATC	TCTCCTAGAG	CACAAGACAA	ATCATAGTTA	CTGGCTATTT
17161	GCAATCTGAA	GCATCTCAAT	ATCCACACCC	TGGGATTAAA	ACAAAAACAT	ATTACATCA
17221	CATAACTGTT	TTTTTTTTTCC	AATTTTTTAT	TAGGTATTTT	CTTTATTTAC	ATTTCAAATG
17281	CTATCCCGAA	AGTCCCCTAT	ACCTCCCAC	CTCCCTGCTC	CCCTACACAC	CCACTCCCAC

Figure 5G:

```
17341 TTTTGTGACCC TGGAGTTCCC CGGTACTGGG GCATATAAAG TTTGCAAGAC CAAGGGGCGCT
17401 CTCTTCCCAG TGATGGCCGA CTAAGCCATC TTCTGCTACA TATGCAGATA GAGACACGAG
17461 CTCTGGGGGT ACTAGTTAGT TCATATTGTT GTTCCACCTA TAGGGTCGCA GACCCCTTCA
17521 GCTCCTTGGG TACTTTGTCT AGCTCCTCCA CTGGGGGGCTC TGTGTTTTAT CTAATAGATG
17581 ACTGTGAGCA TCCACTTCTG TATTTGACAG GCACTGGCCT AGCGTCACAT GAGCCAGCTA
17641 TATCAGGGTC CTTTCAGCAA AACCTTGCTG GCATGTGCAA TAGTGTCTGC GTTTGGTGGT
17701 TGATTATGGG ATGGATCCAC TAGTTCTAGA GCGGCCGCCA CCGCGGTGGA GCTCCAGCTT
17761 TTGTTCCTT TAGTGAGGGT TAATTGCGCG CTTGGCGTAA TCATGGTCAT AGCTGTTTCC
17821 TGTGTGAAAT TGTATCCGC TCACAATTCC ACACAACATA CGAGCCGGAA GCATAAAGTG
17881 TAAAGCCTGG GGTGCCTAAT GAGTGAGCTA ACTCACATTA ATTGCGTTGC GCTCACTGCC
17941 CGCTTTCAG TCGGGAAACC TGTCGTGCCA GCTGCATTAA TGAATCGGCC AACCGCGGG
18001 GAGAGGCGGT TTGCGTATTG GCGCTCTTC CGCTTCCTCG CTCACTGACT CGCTGCGCTC
18061 GGTCGTTCCG CTGCGCGGAG CCGTATCAGC TCACTCAAAG GCGGTAATAC GGTATCCAC
18121 AGAATCAGGG GATAACGCAG GAAAGAACAT GTGAGCAAAA GGCCAGCAAA AGGCCAGGAA
18181 CCGTAAAAAG GCCGCGTTGC TGGCGTTTTT CCATAGGCTC CGCCCCCTG ACGCACTCA
18241 CAAAAATCGA CGCTCAAGTC AGAGGTGGCG AAACCCGACA GGACTATAAA GATACCAGGC
18301 GTTTCCTTCC GGAAGCTCCC TCGTGCGCTC TCCTGTTCCG ACCCTGCCGC TTACCGGATA
18361 CCTGTCCGCC TTTCTCCTT CCGGAAGCGT GCGCTTTCT CATAGCTCAC GCTGAGGTA
18421 TCTCAGTTCG GTGTAGGTCG TTCGCTCCAA GCTGGGCTGT GTGCACGAAC CCCCCGTCA
18481 TGGCCCGCGT TCGCCTTAT TCGCTTACG CCGGTAACTA TCGTCTTGG TCCAACCCGG TAAGACACGA
18541 CTTATCGCCA CTGGCAGCAG CCACTGGTAA CAGGATTAGC AGAGCGAGGT ATGTAGGCGG
18601 TGCTACAGAG TTCTGAAGT GGTGGCCTAA CTACGGCTAC ACTAGAAGGA CAGTATTTGG
18661 TATCTGCGCT CTGCTGAAGC CAGTTACCTT CCGAAAAAGA GTTGGTAGCT CTTGATCCGG
18721 CAAACAAACC ACCGCTGGTA GCGGTGGTTT TTTTGTGTC AAGCAGCAGA TTACGCGCAG
18781 AAAAAAAGGA TCTCAAGAAG ATCCTTTGAT CTTTCTACG GGTCTGACG CTCAGTGGAA
18841 CGAAACTCA CGTTAAGGGA TTTTGGTCAT GAGATTATCA AAAAGGATCT TCACCTAGAT
18901 CCTTTTAAAT TAAAAATGAA GTTTTAAATC AATCTAAAGT ATATATGAGT AAAGTTGGTC
18961 TGACAGTTAC CAATGCTTAA TCAGTGAGGC ACCTATCTCA GCGATCTGTC TATTTCTGTT
19021 ATCCATAGTT GCCTGACTCC CCGTCGTGTA GATAACTACG ATACGGGAGG GCTTACCATC
19081 TGGCCCCAGT GCTGCAATGA TACCGCAGA CCCACGCTCA CCGGCTCCAG ATTTATCAGC
19141 AATAAACAG CCAGCCGGAA GGGCCGAGCG CAGAAGTGGT CCTGCAACTT TATCCGCTC
19201 CATCCAGTCT ATTAATTGTT GCCGGGAAGC TAGAGTAAGT AGTTCGCCAG TTAATAGTTT
19261 GCGCAACGTT GTTGCCATTG CTACAGGCAT CGTGGTGTCA CGCTCGTCGT TTGGTATGGC
19321 TTCATTGAGC TCCGGTCCC AACGATCAAG GCGAGTTACA TGATCCCCCA TGTGTGCAA
19381 AAAAGCGGTT AGTCCTTCG GTCCTCCGAT CGTTGTGAGA AGTAAGTTGG CCGCAGTGTT
19441 ATCACTCATG GTTATGGCAG CACTGCATAA TTCTCTTACT GTCATGCCAT CCGTAAGATG
19501 CTTTCTGTG ACTGGTGAGT ACTCAACCAA GTCATTCTGA GAATAGTGTA TGCGCGGACC
19561 GAGTTGCTCT TGCCCCGCGT CAATACGGGA TAATAACGCG CCACATAGCA GAACCTTAAA
19621 AGTGCTCATC ATTGAAAAAC GTTCTTCGGG GCGAAAACTC TCAAGGATCT TACCGCTGTT
19681 GAGATCCAGT TCGATGTAAC CCACTCGTGC ACCCAACTGA TCTTCAGCAT CTTTACTTTT
19741 CACCAGCGTT TCTGGGTGAG CAAAAACAGG AAGGCAAAAT GCCGCAAAA AGGGAATAAG
19801 GCGACACGGA AAATGTTGAA TACTCATACT CTTCCTTTTT CAATATTATT GAAGCATTTA
19861 TCAGGGTTAT TGTCTCATGA GCGGATACAT ATTTGAATGT ATTTAGAAAA ATAAACAAAT
19921 AGGGGTTCCG CGCACATTTT CCGGAAAAGT GCCACCTAAA TTGTAAGCGT TAATATTTTG
19981 TTAAAAATTC CGTTAAATTT TTGTTAAATC AGCTCATTTT TTAACCAATA GGCCGAAATC
20041 GGC AAAATCC CTTATAAATC AAAAGAATAG ACCGAGATAG GGTGAGTGT TGTTCAGTT
20101 TGAACAAGA GTCCACTATT AAAGAACGTG GACTCCAACG TCAAAGGGCG AAAAACCGTC
20161 TATCAGGGCG ATGGCCCACT ACGTGAACCA TCACCCTAAT CAAGTTTTTT GGGGTCGAGG
20221 TGCCGTAAAG CACTAAATCG GAACCCTAAA GGGAGCCCC GATTTAGAGC TTGACGGGGA
20281 AAGCCGCGGA ACGTGGCGAG AAAGGAAGGG AAGAAAGCGA AAGGAGCGGG CGTAGGGCG
20341 CTGGCAAGTG TAGCGGTCAC GCTGCGCGTA ACCACCACAC CCGCCGCGCT TAATGCGCCG
20401 CTACAGGGCG CGTCCCATTC GCCATTACAG CTGCGCAACT GTTGGGAAGG GCGATCGGTG
20461 CGGGCTCTT CGCTATTACG CCAGCTGGCG AAAGGGGGAT GTGCTGCAAG GCGATTAAGT
20521 TGGGTAACGC CAGGGTTTTT CCAATCACGA CGTTGTAATA CGACGGCCAG TGAGCGCGCG
20581 TAATACGACT CACTATAGGG CGAATTGGGT ACCGGGCCCC CCC
```

//

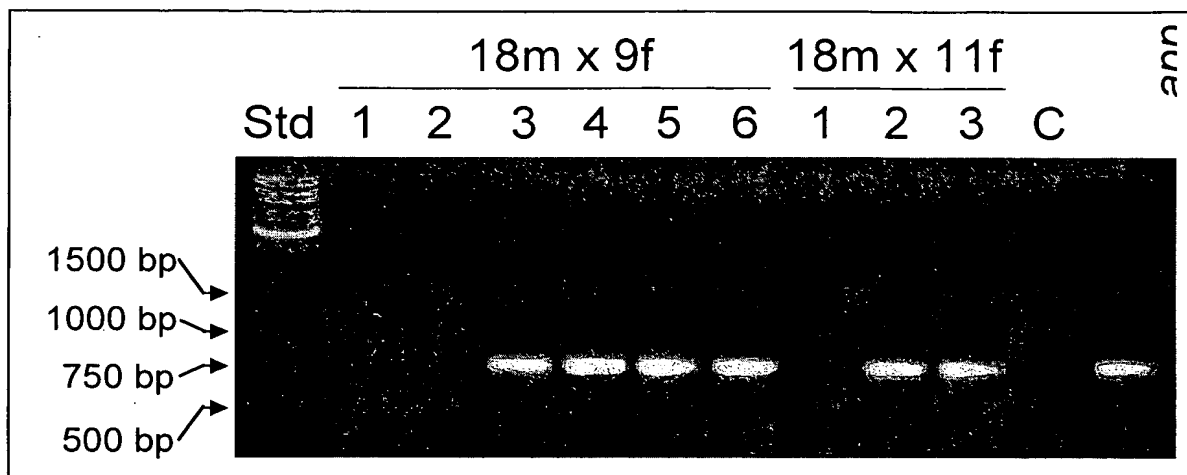


Figure 6

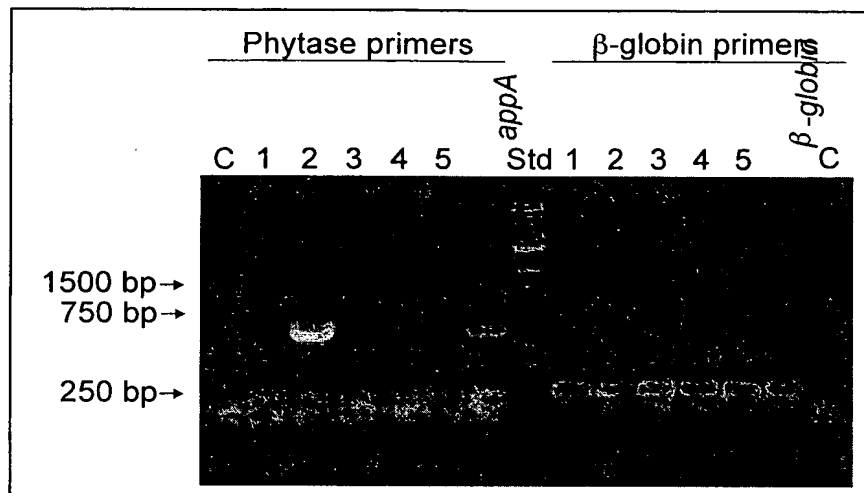


Figure 7

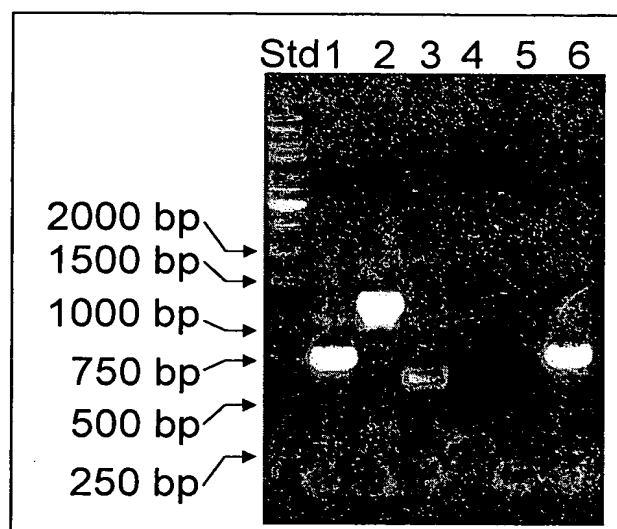


Figure 8

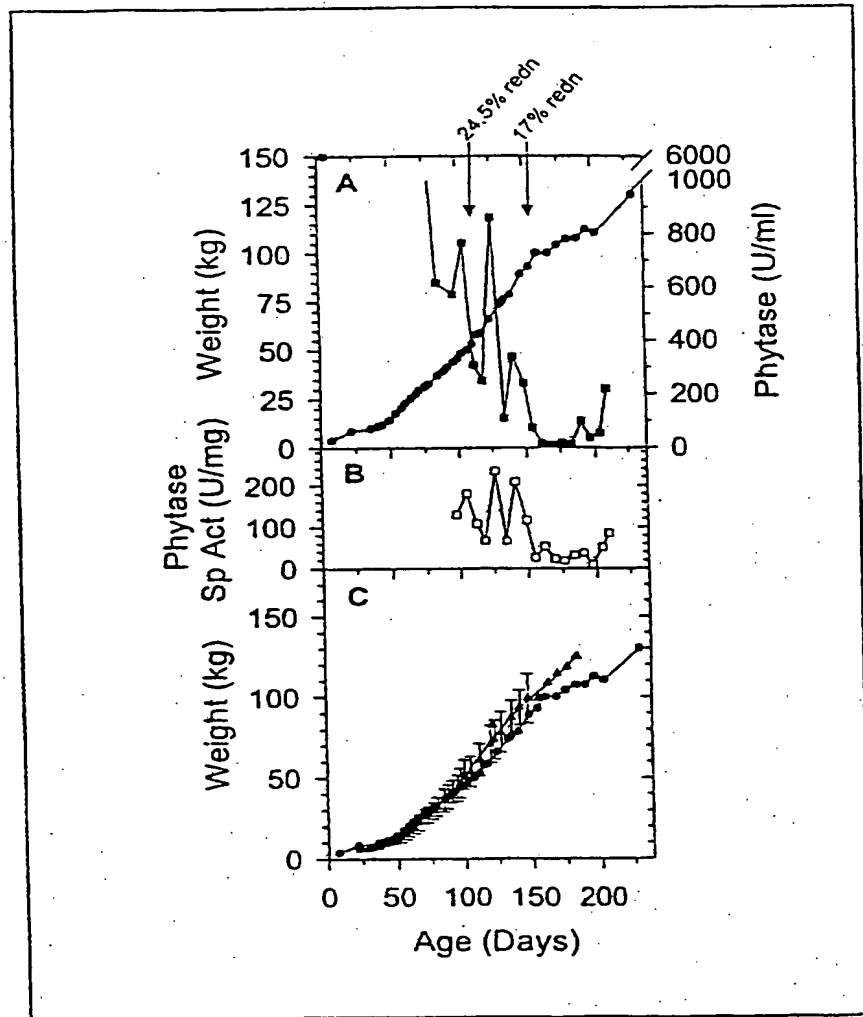


Figure 9

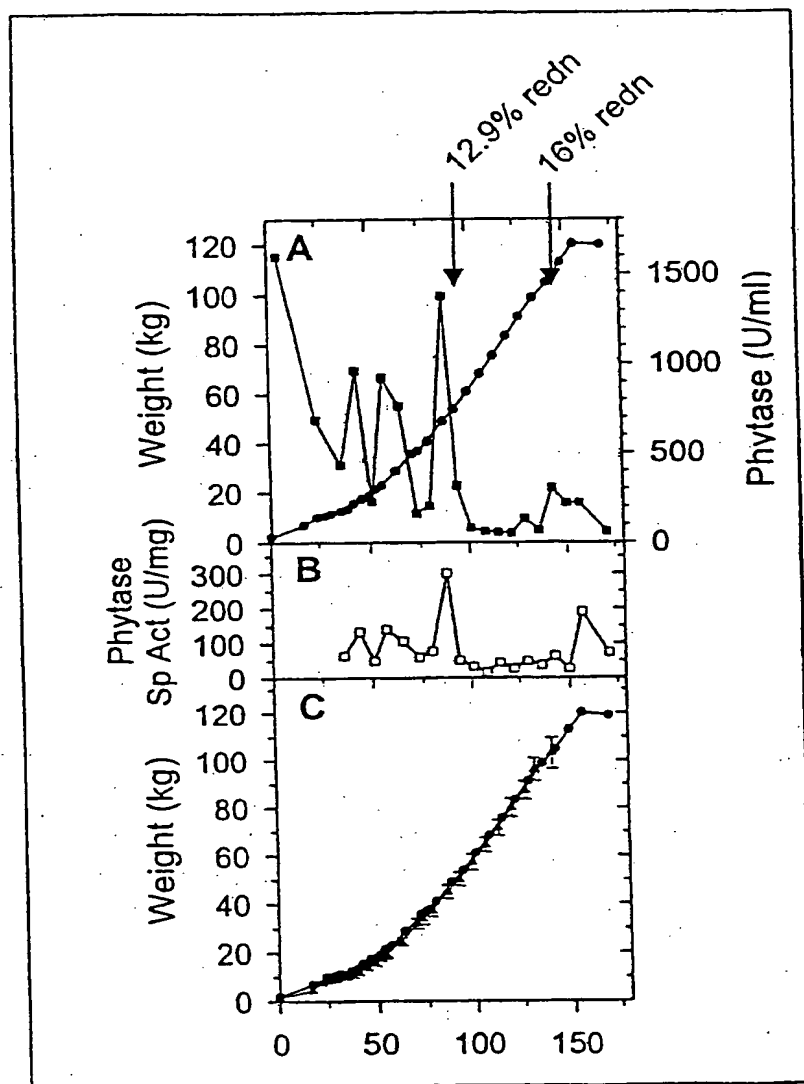


Figure 10

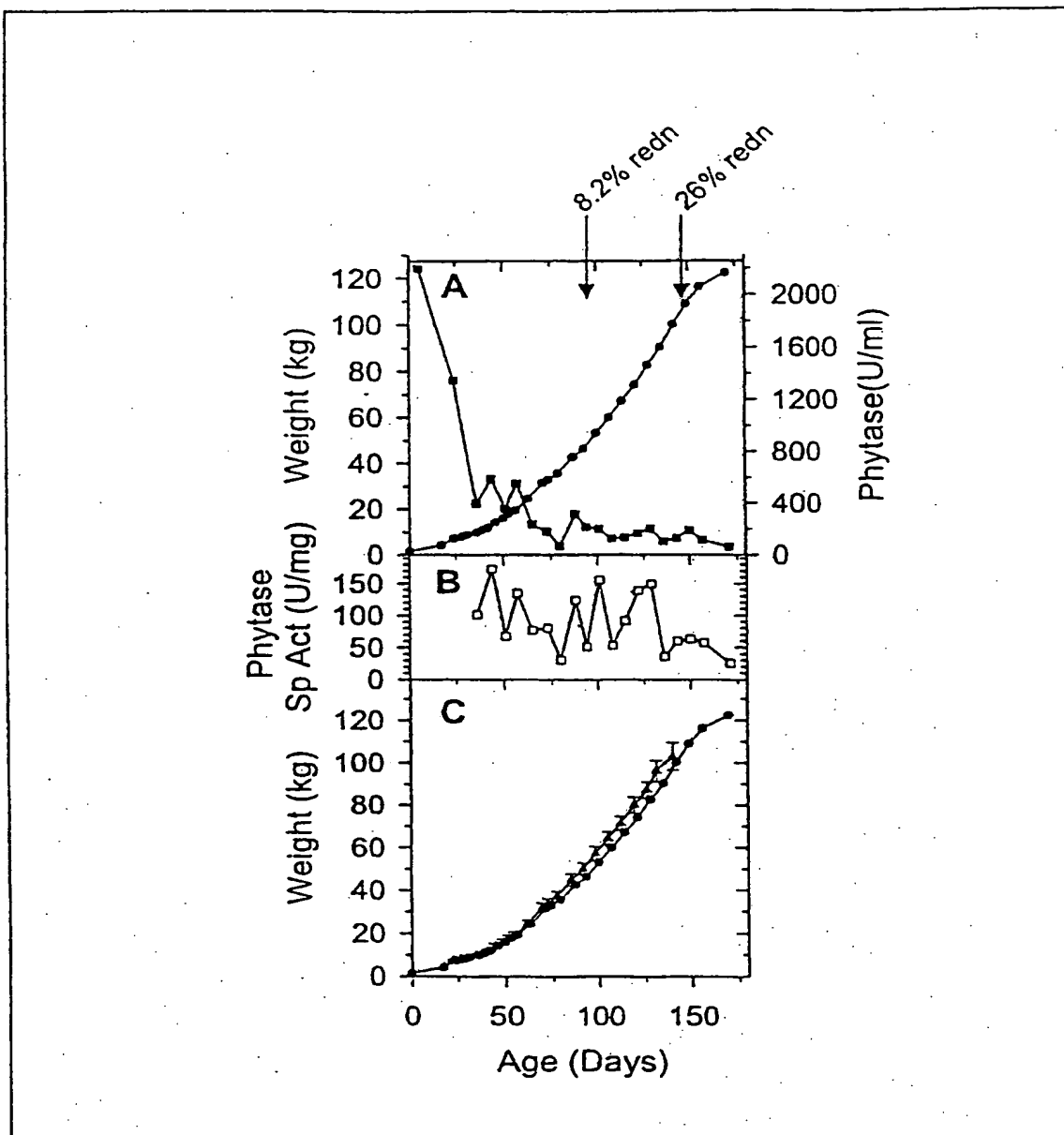


Figure 11

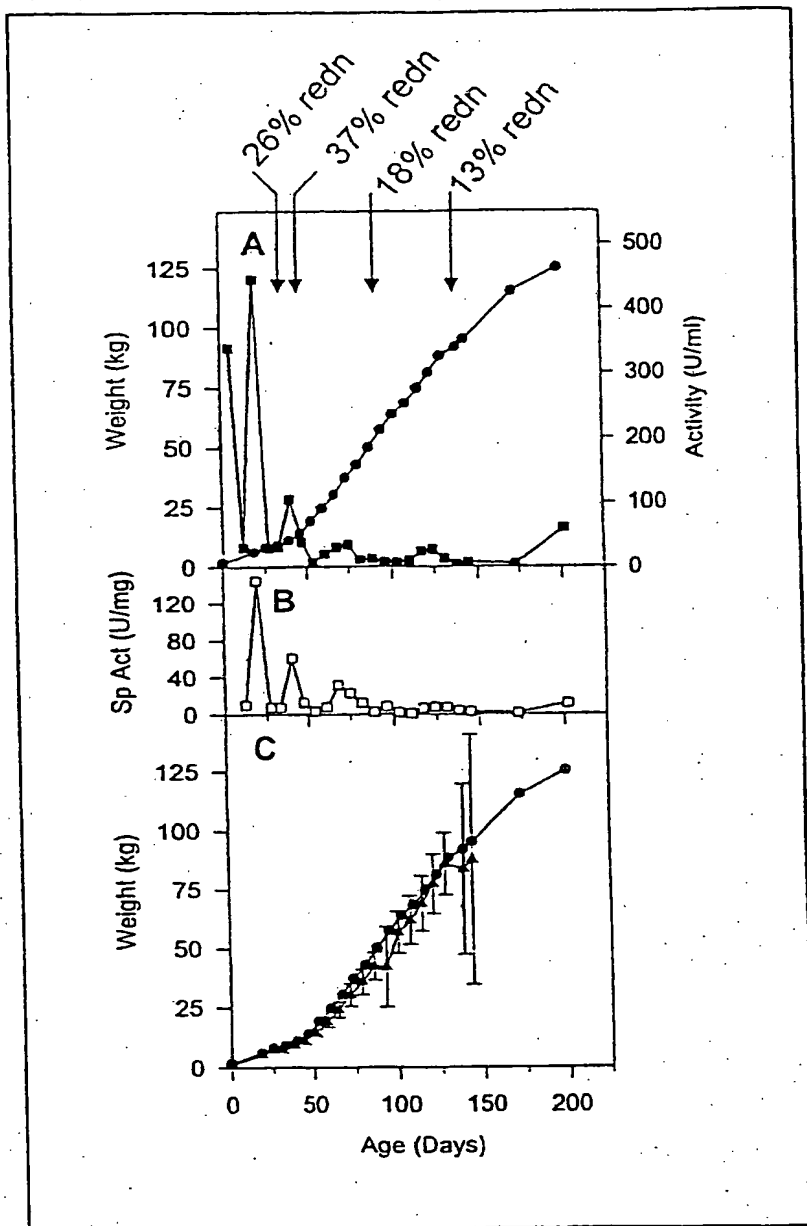


Figure 12

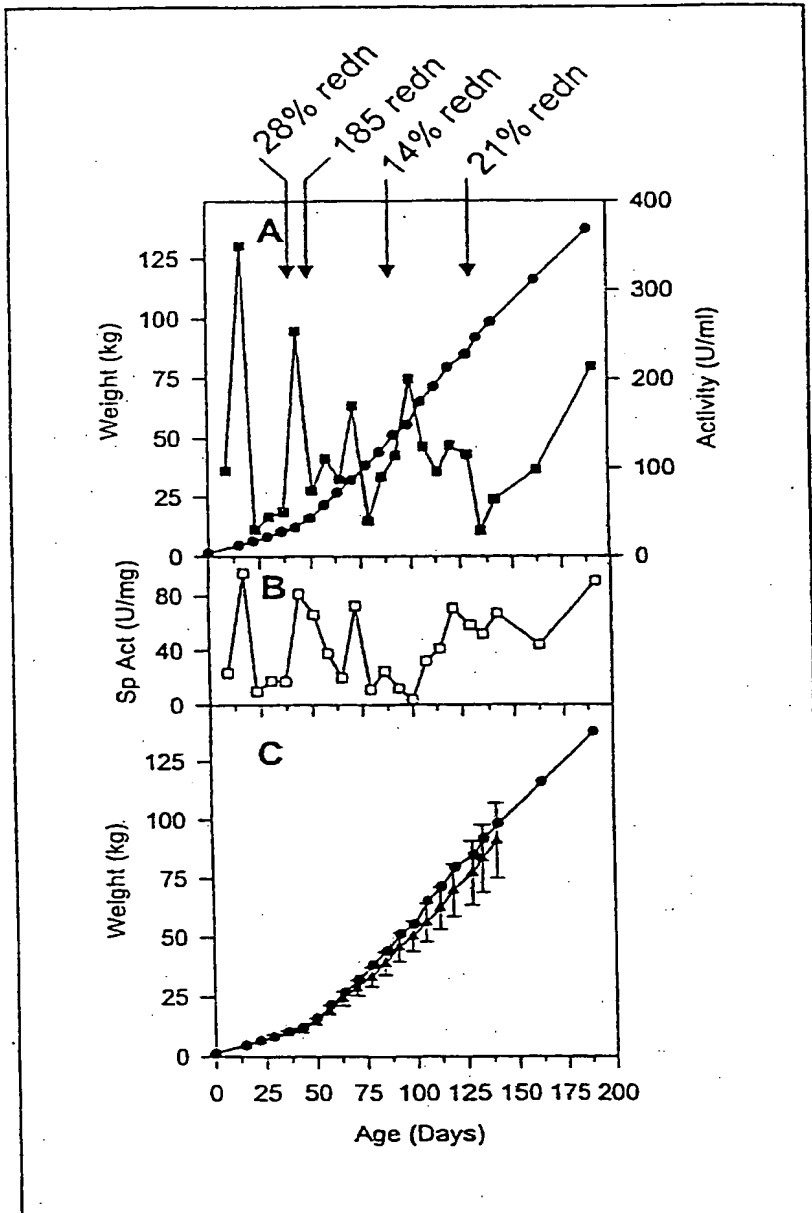


Figure 13

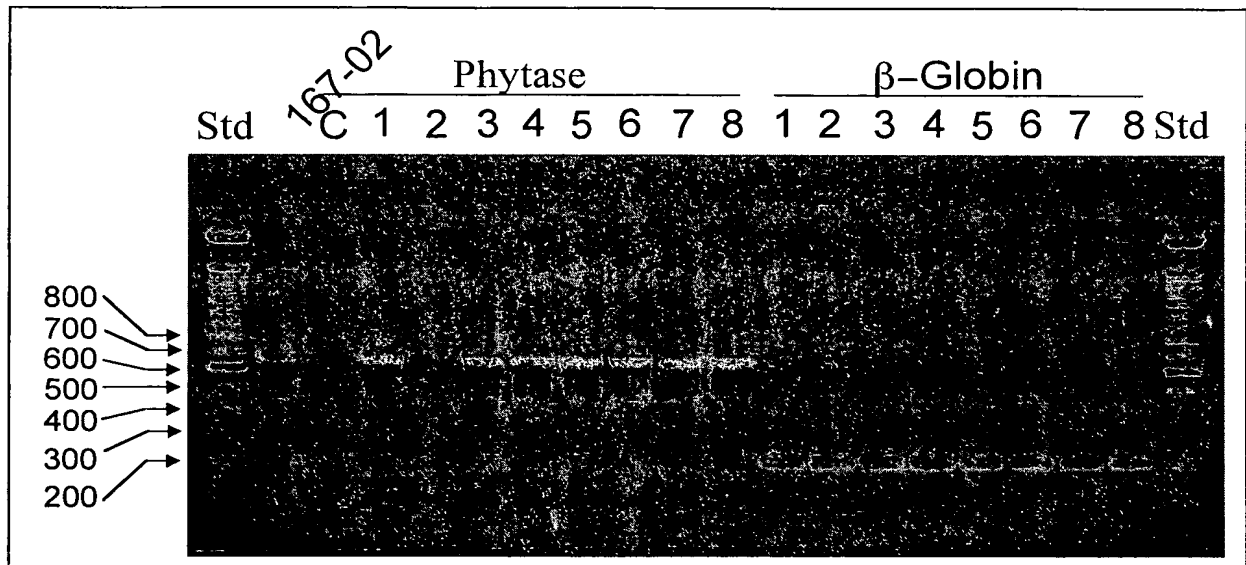


Figure 14

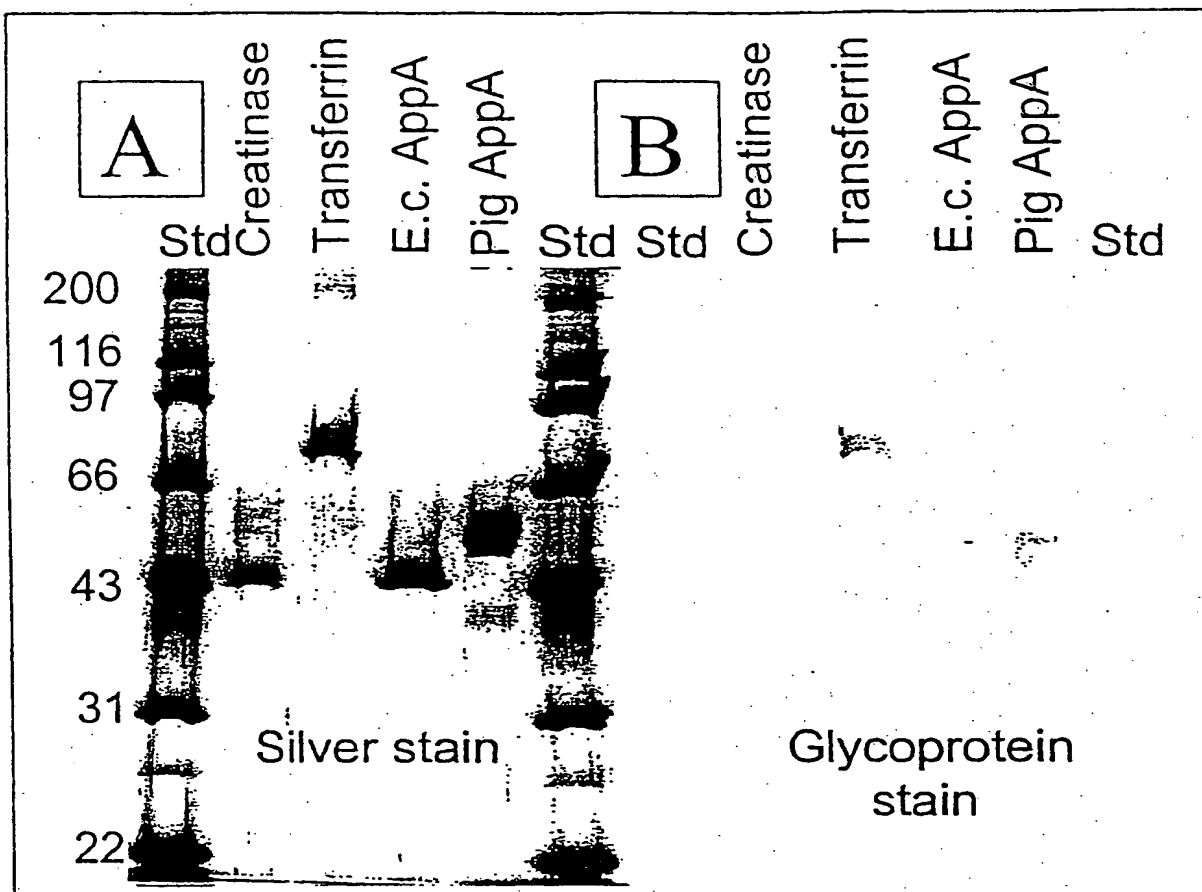


Figure 15

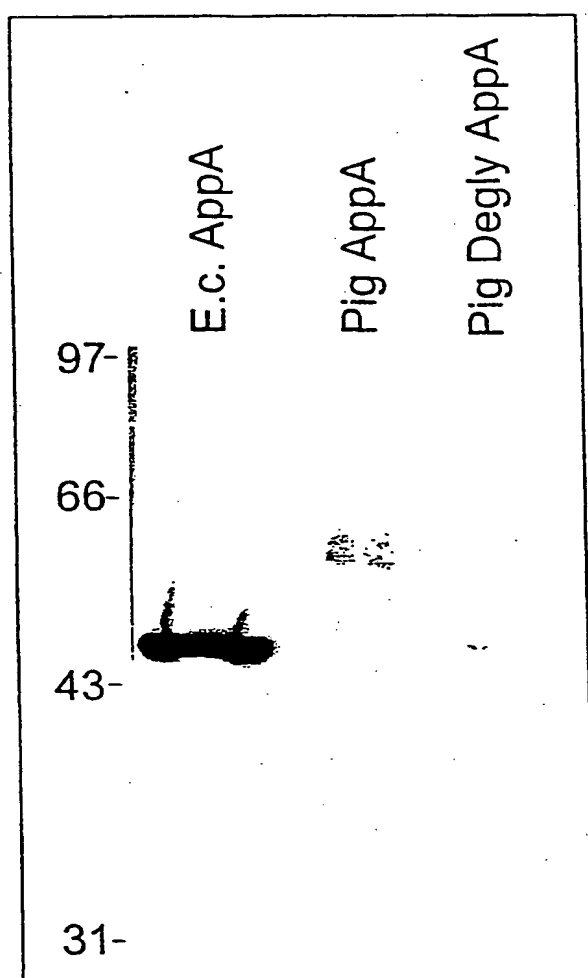


Figure 15B

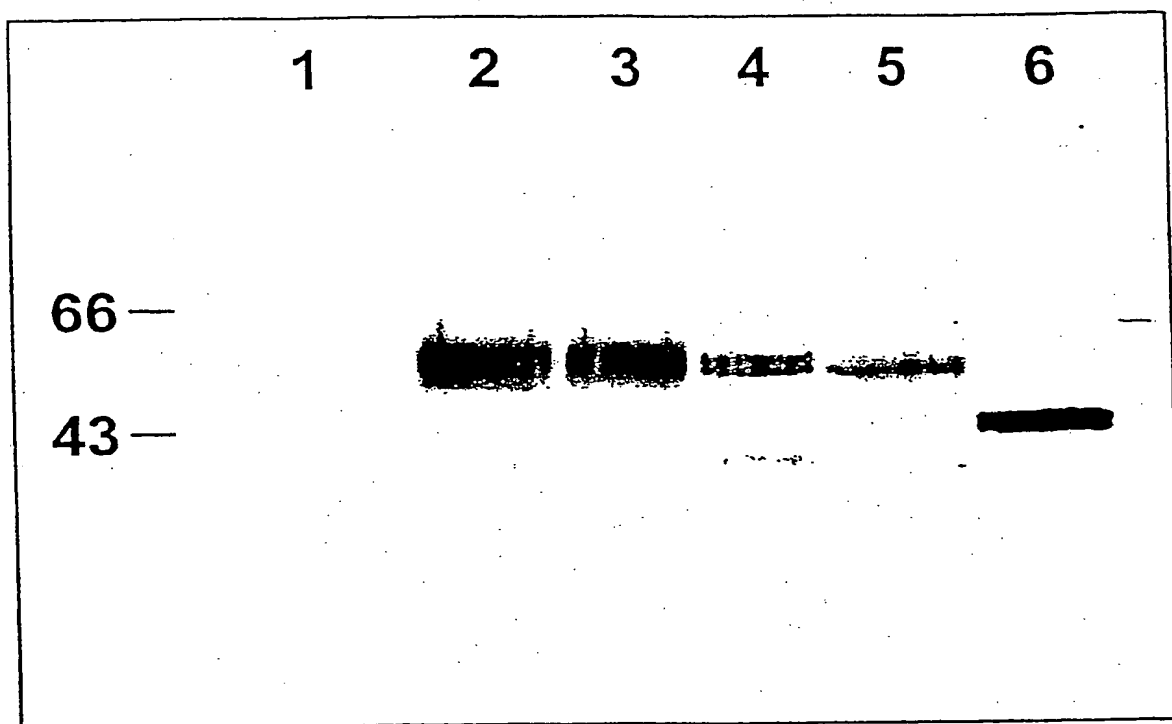


Figure 16

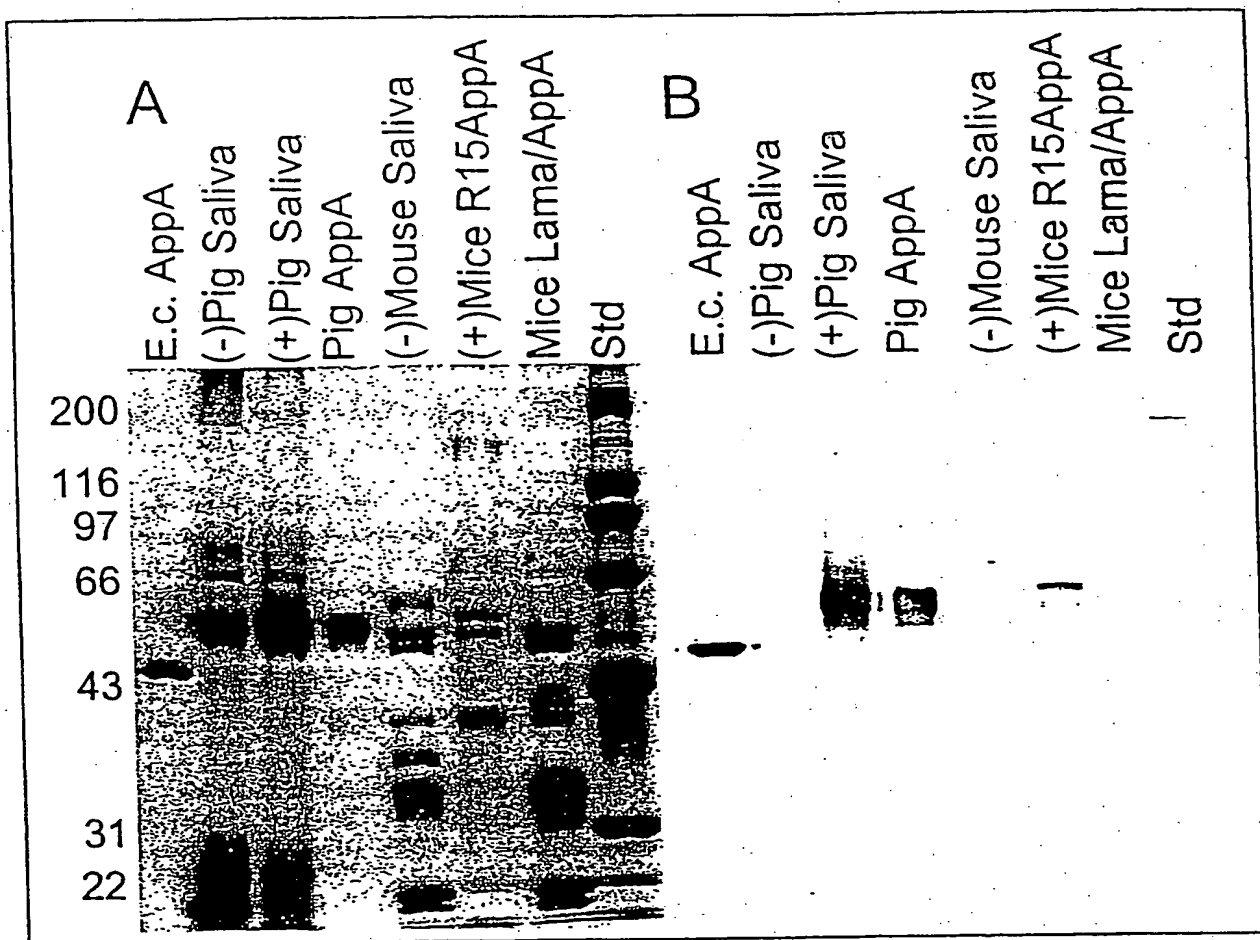


Figure 17

Figure 18: Nucleic acid sequence of the known segment of the R15/appa+intron plasmid, including the vector sequences of pBLCAT3 (SEQ ID NO:2).

LOCUS R15/appa+intron 6708 bp DNA SYN 15-APR-2000
 DEFINITION R15/appa+intron transgene with vector cut 13543 to 4954
 ACCESSION R15/appa+intron
 REFERENCE 1 (bases 1 to 6708))
 SOURCE synthetic construct.
 ORGANISM synthetic construct
 artificial sequence.
 KEYWORDS salivary proline-rich protein, acid glucose-1-phosphatase; appA
 gene; periplasmic phosphoanhydride phosphohydrolase; artificial
 sequence;
 AUTHORS Golovan, S., Forsberg, C.W., Phillips, J.
 JOURNAL Unpublished.

DEFINITION Rat salivary proline-rich protein (RP15) gene.
 ACCESSION M64793 M36414
 VERSION M64793.1 GI:206711
 SOURCE Rat (Sprague-Dawley) liver DNA.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
 Mammalia;
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 1748)
 AUTHORS Lin, H.H. and Ann, D.K.
 TITLE Molecular characterization of rat multigene family
 encoding
 proline-rich proteins
 JOURNAL Genomics 10, 102-113 (1991)
 MEDLINE 91257817
 FEATURES Location/Qualifiers
 source 1..1748
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /tissue_type="liver"
 /tissue_lib="cosmid genomic library"
 misc_feature 1802-1810
 /function=" consensus sequence for initiation in
 higher eukaryotes "

FEATURES Location/Qualifiers
 DEFINITION E. coli periplasmic phosphoanhydride phosphohydrolase (appa)
 gene,

ACCESSION M58708 L03370 L03371 L03372 L03373 L03374 L03375
 VERSION M58708.1 GI:145283
 SOURCE Escherichia coli DNA.
 ORGANISM Escherichia coli
 Bacteria; Proteobacteria; gamma subdivision;
 Enterobacteriaceae;
 Escherichia.
 REFERENCE 1 (bases 1811..3109)
 AUTHORS Dassa, J., Marck, C. and Boquet, P.L.

Figure 18A:

```
TITLE      The complete nucleotide sequence of the Escherichia coli
            gene appA reveals significant homology between pH 2.5
            acid phosphatase and glucose-1-phosphatase
JOURNAL    J. Bacteriol. 172 (9), 5497-5500 (1990)
MEDLINE    90368616

FEATURES   Location/Qualifiers
Source          1811..3109
                /organism="Escherichia coli"
                /db_xref="taxon:562"
sig_peptide    1811.. 1876
                /gene="appA"
CDS            1811..3109
                /gene="appA"
                /standard_name="acid phosphatase/phytase"
                /transl_table=11
                /product="periplasmic phosphoanhydride
                phosphohydrolase"
                /protein_id="AAA72086.1"
                /db_xref="GI:145285"

mat_peptide    1877  3106
                /gene="appA"
                /product="periplasmic phosphoanhydride
                phosphohydrolase"

mutation       replace(1817.. 1819,"gcg changed to gcc")
                /gene="appA"
                /standard_name="A3 mutant"
                /note="created by site directed mutagenesis"
                /phenotype="silent mutation"
mutation       replace(3092..3094," ccg changed to ccc")
                /gene="appA"
                /standard_name=" P428 mutant"
                /note="created by site directed mutagenesis"
                /phenotype=" silent mutation "
mutation       replace(3095..3097," gcg changed to gct")
                /gene="appA"
                /standard_name=" A429 mutant"
                /note="created by site directed mutagenesis"
                /phenotype=" silent mutation "
```

Figure 18B:

DEFINITION Plasmid pBLCAT3 (bases 3109 to 6708)

ACCESSION X64409

VERSION X64409.1 GI:58163

SOURCE synthetic construct.

ORGANISM synthetic construct
artificial sequence.

REFERENCE 1 (bases 3109 to 6708)

AUTHORS Luckow,B.H.R.

TITLE Direct Submission

JOURNAL Submitted (06-FEB-1992) B.H.R. Luckow, German Cancer Res
Center, Im Neuenheimer Feld 280, W-6900 Heidelberg, FRG

REFERENCE 2 (bases 3109 to 6708)

AUTHORS Luckow,B. and Schutz,G.

TITLE CAT constructions with multiple unique restriction sites

for the functional analysis of eukaryotic promoters and
regulatory elements

JOURNAL Nucleic Acids Res. 15 (13), 5490 (1987)

MEDLINE 87260024

COMMENT Promoterless CAT vector for transient transfection
experiments with eukaryotic cells. Allows the analysis of foreign
promoters and enhancers.

FEATURES Location/Qualifiers

source 3109 to 6116
/organism="synthetic construct"
/db_xref="taxon:32630"

SV40 t intron 3197..3810
/note="SV40 signals"

polyA_signal 3807..4047
/note="SV40 signals"

CDS complement(5244..6104)
/codon_start=1
/transl_table=11
/gene="Amp"
/product="beta-lactamase"
/protein_id="CAA45753.1"
/db_xref="GI:58165"

BASE COUNT 1916 a 1479 c 1515 g 1798 t

ORIGIN

1 GGATCCCCTT TGCTATGTAG TTTTAAATGG AAATTACAAC CCATAGTGTG TTGATAAATA
61 GAGAGTCCTG TTTGGTTTAA GCAACCTCTG TTTCTCATAA ACTCCATAAA AACAGGAATA
121 CTCTTTGTTT CTAGCATAAC CAAAAGATTT AGTGAATTGA AAACAATGTT CCCTTAGAGT
181 ATAGGTCTAA TAACCCCGAA AATATTACCA TGATACTGAG CATTTGTAAG TATCTCATAG
241 CATGTAGTAT CCATAGTCCA TCAATGAGAG AGACATTTAA CATGATTTTC ATTAATCAGG
301 TGGAAAAGAC ATGACAACAT TCACAGGCAC TGCACAGAAC ATAGTGGTCC ACCTTGCACA
361 TATTTCACTA AACTAGGTTT ATCTATTTTG TTGCTTTCTC TAACATCTCT GCAATGAAGC
421 AGGTCAACAG TGCCACATAT CCTTTACTTA ACCTAAGGAA CACAAAAAAT TTTCTACATA
481 TATCTGGTGT AGAGAGTGCT TAAAATAAGT TTTCCAAGAA TGGAAAAGAA ATGTTCTGAC
541 TTAACAATTA AGACAGTATT TATTTAAAGC AAGAAATATG AGGCACACAA GAAAATATTT
601 TGGGAAGAAA CCATTTGGTG AACAATATTT CAAATAAAAA TAGACAAACA TAGTTAATTG
661 TAAAACATAT GTTTGACCAG CCCTTCTTTT CAATAGGCTT AATGTGAATA AAATGTTAAA
721 GATTCTCTTT GGGTGGCTGC AAATTGTCCA CGAATAAGAC AAAATATAAA AATAAGGACT
781 GAGTCTCACA AAATGAAAAG GAAATATATT CAGAAAAGAGA ATCTTGAGAG AATGTGTTGT
841 CACAAATTAA AGAAAACCTG TGGTGAATGA CATCCTGAGG CCTGAGCTAT TACTGACATT

Figure 18C:

```
901 TAAGATAAAG GTAACGTGTAT ACATTTGTCC CATTGAGGGG ACAAGAAAGC TGCTCTCATG
961 TTCAGCTCTA TAATTCTTGC CTTAAACAAC TTAAATAGAA TGATTTAAAA TATGGAGCTG
1021 TCCATGGACC TTTGAAATAT AAAATAGTCA AGCAACTTAT CAAGGAATTA CAGATTCCTT
1081 GATACTAACA CAGGTAAATC CCACACGTGT TTTCACTACTA CATTGCTGGG GATTTTATTG
1141 ATGTAATAGG TCACATGTTT TTCGGGCCAA TGTTGCTGTT ATTCGGTTAC TTCAAGAGAA
1201 TAGTGGCAAC TGATGCTATG TATTCTAGGG GTTTGAAGTG ATGTTTCATG ATTGAAATTT
1261 GTAAAAGAAT AACATCATCA TTCTTAACAA TAGAACATAT AAAGTCACAC AGAAGTGACA
1321 GTGTTTAAAG TGTACTATTG ATCAAAGAAA TTTATTACCT TCAGTTTCAA TGGAAATAAT
1381 TACTGATAAT ACAAACATGT GTGAACACAC ACTAATCCTA TCCAAATGCA CAGTGATACA
1441 CAGAAAATAT TAGCAAGTAG AATGCAATAT TTATATAACG ATTGTATTTA TCAATCAATT
1501 GTATGTATCA ATATATGGGC TATTTTCTTA CACATGATTT TATTCAAATT TACTCTAATC
1561 ATTGTTGAAC CATTTAGAAA AGGCATACTG GCAACTTTTC CTTACCTCAT CCAGCTGGGC
1621 AAAAGTCCCA GTGTGGAGTA AAGGATGCAA GATTTCTCTG TCTGTTAAGT ATAAAATAAT
1681 AGTATGAATT CAAAGGTGCC ATTCTTCTGC TTCTAGTTAT AAAGGCAGTG CTTGCTTCTT
1741 CCAGCACAGA TCTGGATCTC GAGGAGCTTG GCGAGATTTT CAGGAGCTAA GGAAGCTAAA
1801 AGCCGCCACC ATGAAAGCCA TCTTAATCCC ATTTTATCTT CTTCTGATTC CGTTAACCCC
1861 GCAATCTGCA TTCGCTCAGA GTGAGCCGGA GCTGAAGCTG GAAAGTGTGG TGATTGTCAG
1921 TCGTCATGGT GTGCGTGCTC CAACCAAGGC CACGCAACTG ATGCAGGATG TCACCCCAGA
1981 CGCATGGCCA ACCTGGCCGG TAAACCTGGG TTGGCTGACA CCGCGCGGTG GTGAGCTAAT
2041 CGCCTATCTC GGACATTACC AACGCCAGCG TCTGGTAGCC GACGGATTGC TGGCGAAAAA
2101 GGGCTGCCCC CAGTCTGGTC AGGTCGCGAT TATTGCTGAT GTCGACGAGC GTACCCGTAA
2161 AACAGGCGAA GCCTTCGCCG CCGGGCTGGC ACCTGACTGT GCAATAACCG TACATACCCA
2221 GGCAGATACG TCCAGTCCCG ATCCGTTATT TAATCCTCTA AAAACTGGCG TTTGCCAACT
2281 GGATAACGCG AACGTGACTG ACGCGATCCT CAGCAGGGCA GGAGGGTCAA TTGCTGACTT
2341 TACCGGGCAT CGGCAAACGG CGTTTCGCGA ACTGGAACGG GTGCTTAATT TTCCGCAATC
2401 AAACCTGTGC CTTAAACGTG AGAAACAGGA CGAAAGCTGT TCATTAACGC AGGCATTACC
2461 ATCGGAACCT AAGGTGAGCG CCGACAATGT CTCATTAACC GGTGCGGTAA GCCTCGCATC
2521 AATGCTGACG GAGATATTTT TCCTGCAACA AGCACAGGGA ATGCCGGAGC CGGGGTGGGG
2581 AAGGATCACC GATTACACAC AGTGGAACAC CTTGCTAAGT TTGCATAACG CGCAATTTTA
2641 TTTGCTACAA CGCACGCCAG AGGTTGCCCG CAGCCGCGCC ACCCGTTAT TAGATTTGAT
2701 CAAGACAGCG TTGACGCCCC ATCCACCGCA AAAACAGGCG TATGGTGTGA CATTACCCAC
2761 TTCAGTGCTG TTTATCGCCG GACACGATAC TAATCTGGCA AATCTCGGCG GCGCACTGGA
2821 GCTCAACTGG ACGCTTCCCG GTCAGCCGGA TAACACGCCG CCAGGTGGTG AACTGGTGTG
2881 TGAACGCTGG CGTCGGCTAA GCGATAACAG CCAGTGGATT CAGGTTTCGC TGGTCTTCCA
2941 GACTTTACAG CAGATGCGTG ATAAAACGCC GCTGTCATTA AATACGCCGC CCGGAGAGGT
3001 GAAACTGACC CTGGCAGGAT GTGAAGAGCG AAATGCGCAG GGCATGTGTT CGTTGGCAGG
3061 TTTTACGCAA ATCGTGAATG AAGCACGCAT ACCCGCTTGC AGTTTGTAA GGTATAAGGCA
3121 GTTATTGGTG CCCTTAAACG CCTGGTGCTA CGCCTGAATA AGTGATAATA AGCGGATGAA
3181 TGGCAGAAAT TCGCCGGATC TTTGTGAAGG AACCTTACTT CTGTGGTGTG ACATAATTGG
3241 ACAAACCTACC TACAGAGATT TAAAGCTCTA AGGTAAATAT AAAATTTTTT AGTGATATAAT
3301 GTGTTAAACT ACTGATTCTA ATTGTTTGTG TATTTTAGAT TCCAACCTAT GGAAGTATG
3361 AATGGGAGCA GTGGTGAAT GCCTTTAATG AGGAAAACCT GTTTTGCTCA GAAGAAATGC
3421 CATCTAGTGA TGATGAGGCT ACTGCTGACT CTCAACATTC TACTCCTCCA AAAAAGAAGA
3481 GAAAGGTAGA AGACCCCAAG GACTTTCCTT CAGAATTGCT AAGTTTTTTT AGTCATGCTG
3541 TGTTTAGTAA TAGAACTCTT GCTTGCTTTG CTATTTACAC CACAAAGGAA AAAGCTGCAC
3601 TGCTATACAA GAAAATTATG GAAAAATATT CTGTAACCTT TATAAGTAGG CATAACAGTT
3661 ATAATCATAA CATACTGTTT TTTCTTACTC CACACAGGCA TAGAGTGTCT GCTATTAATA
3721 ACTATGCTCA AAAATTGTGT ACCTTTAGCT TTTTAATTTG TAAAGGGGTT AATAAGGAAT
3781 ATTTGATGTA TAGTGCCTTG ACTAGAGATC ATAATCAGCC ATACCACATT TGTAGAGGTT
3841 TTAATTGCTT TAAAAACCT CCCACACCTC CCCCTGAACC TGAAACATAA AATGAATGCA
3901 ATGTTGTTG TTAACCTGTT TATTGCAGCT TATAATGGTT ACAAATAAAG CAATAGCATC
3961 ACAAATTTCA CAAATAAAGC ATTTTTTTTCA CTGCATTCTA GTTGTGGTTT GTCCAAACTC
4021 ATCAATGTAT CTTATCATGT CTGGATCGAT CCCCCGGTAC CGAGCTCGAA TTCGTAATCA
4081 TGGTCATAGC TGTTTCCTGT TGAAATTTGT TATCCGCTCA CAATTCCACA CAACATACGA
4141 GCCGGAAGCA TAAAGTGTA AGCCTGGGTT GCCTAATGAG TGAGCTAAGT CACATTAATT
4201 GCGTTGCGCT CACTGCCCGC TTTCCAGTCG GGAAACCTGT CGTGCCAGCT CGATTAATGA
4261 ATCGGCCAAC GCGCGGGGAG AGGCGGTTTG CGTATTGGGC GCTCTTCCGC TTCTCGCTC
4321 ACTGACTCGC TGCGCTCGGT CGTTCGGCTG CCGCGAGCGG TATCAGCTCA CTCAAAGGCG
```

Figure 18D:

```
4381 GTAATACGGT TATCCACAGA ATCAGGGGAT AACGCAGGAA AGAACATGTG AGCAAAAGGC
4441 CAGCAAAAGG CCAGGAACCG TAAAAAGGCC GCGTTGCTGG CGTTTTTCCA TAGGCTCCGC
4501 CCCCCTGACG AGCATCACAA AAATCGACGC TCAAGTCAGA GGTGGCGAAA CCCGACAGGA
4561 CTATAAAGAT ACCAGGCGTT TCCCCCTCGA AGCTCCCTCG TGCCTCTCC TGTTCCGACC
4621 CTGCCGCTTA CCGGATACCT GTCCGCCTTT CTCCCTTCGG GAAGCGTGGC GCTTTCTCAA
4681 TGCTCACGCT GTAGGTATCT CAGTTCGGTG TAGGTCGTTT GCTCCAAGCT GGGCTGTGTG
4741 CACGAACCCC CCGTTCAGCC CGACCGCTGC GCCTTATCCG GTAACATATCG TCTTGAGTCC
4801 AACCCGGTAA GACACGACTT ATCGCCACTG GCAGCAGCCA CTGGTAACAG GATTAGCAGA
4861 GCGAGGTATG TAGGCGGTGC TACAGAGTTC TTGAAGTGGT GGCCTAACTA CGGCTACACT
4921 AGAAGGACAG TATTTGGTAT CTGCGCTCTG CTGAAGCCAG TTACCTTCGG AAAAAAGAGTT
4981 GGTAGCTCTT GATCCGGCAA ACAAACCACC GCTGGTAGCG GTGGTTTTTT TGTTTGCAAG
5041 CAGCAGATTA CGCGCAGAAA AAAAGGATCT CAAGAAGATC CTTTGATCTT TTCTACGGGG
5101 TCTGACGCTC AGTGGAACGA AAATCACGT TAAGGGATTT TGGTCATGAG ATTACAAAA
5161 AGGATCTTCA CCTAGATCCT TTTAAATTAA AAATGAAGTT TTAATCAAT CTAAAGTATA
5221 TATGAGTAAA CTTGGTCTGA CAGTTACCAA TGCTTAATCA GTGAGGCACC TATCTCAGCG
5281 ATCTGTCTAT TTCGTTTCATC CATAGTTGCC TGACTCCCCG TCGTGTAGAT AACTACGATA
5341 CGGGAGGGCT TACCATCTGG CCCCAGTGCT GCAATGATAC CGCGAGACCC ACGCTCACCG
5401 GCTCCAGATT TATCAGCAAT AAACCAGCCA GCCGGAAGGG CCGAGCGCAG AAGTGGTCCCT
5461 GCAACTTTAT CCGCCTCCAT CCAGTCTATT AATTGTTGCC GGGAAGCTAG AGTAAGTAGT
5521 TCGCCAGTTA ATAGTTTGCG CAACGTTGTT GCCATTGCTA CAGGCATCGT GGTGTCACGC
5581 TCGTCGTTTG GTATGGCTTC ATTCAGCTCC GGTTCCTAAC GATCAAGGCG AGTTACATGA
5641 TCCCCCATGT TGTGCAAAAA AGCGGTTAGC TCCTTCGGTC CTCCGATCGT TGTGAGAAGT
5701 AAGTTGGCCG CAGTGTTATC ACTCATGGTT ATGGCAGCAC TGCATAATTC TCTTACTGTC
5761 ATGCCATCCG TAAGATGCTT TTCTGTGACT GGTGAGTACT CAACCAAGTC ATTCTGAGAA
5821 TAGTGTATGC GGCGACCGAG TTGCTCTTGC CCGGCGTCAA TACGGGATAA TACCGCGCCA
5881 CATAGCAGAA CTTTAAAAGT GCTCATCATT GGAAAACGTT CTTCGGGGCG AAAACTCTCA
5941 AGGATCTTAC CGCTGTTGAG ATCCAGTTCG ATGTAACCCA CTCGTGCACC CAACTGATCT
6001 TCAGCATCTT TTACTTTTAC CAGCGTTTCT GGGTGAGCAA AAACAGGAAG GCAAAATGCC
6061 GCAAAAAAGG GAATAAGGGC GACACGAAA TGTGAATAC TCATACTCTT CCTTTTCAA
6121 TATTATTGAA GCATTTATCA GGGTTATTGT CTCATGAGCG GATACATATT TGAATGTATT
6181 TAGAAAAATA AACAAATAGG GGTTCGCGC ACATTTCCCC GAAAAGTGCC ACCTGACGTC
6241 TAAGAAACCA TTATTATCAT GACATTAACC TATAAAAATA GGCGTATCAC GAGGCCCTTT
6301 CGTCTCGCGC GTTTCGGTGA TGACGGTGAA AACCTCTGAC ACATGCAGCT CCCGAGACG
6361 GTCACAGCTT GTCTGTAAGC GGATGCCGGG AGCAGACAAG CCCGTCAGGG CGCGTCAGCG
6421 GGTGTTGGCG GGTGTCGGGG CTGGCTTAAC TATGCGGCAT CAGAGCAGAT TGTACTGAGA
6481 GTGCACCATA TGCGGTGTGA AATACCGCAC AGATGCGTAA GGAGAAAATA CCGCATCAGG
6541 CGCCATTTCG CATTCAGGCT GCGCAACTGT TGGGAAGGGC GATCGGTGCG GGCCTCTTCG
6601 CTATTACGCC AGCTGGCGAA AGGGGGATGT GCTGCAAGGC GATTAAGTTG GGTAAACCCA
6661 GGGTTTTCCC AGTCACGACG TTGTAAAACG ACGGCCAGTG CCAAGCTT
```

//

Figure 19: Nucleic acid sequence of the known segment of the R15/appa+intron transgene used for the generation of transgenic mice (SEQ ID NO: 3).

```

LOCUS      R15/appa      4060 bp      DNA      SYN      15-APR-2000
DEFINITION R15/appa transgene without vector
ACCESSION  R15/appa
REFERENCE  1  (bases 1 to 4060)
SOURCE     synthetic construct.
           ORGANISM  synthetic construct
                     artificial sequence.
KEYWORDS   salivary proline-rich protein, acid glucose-1-phosphatase; appA
           gene; periplasmic phosphoanhydride phosphohydrolase; artificial
           sequence;
AUTHORS    Golovan, S., Forsberg, C.W., Phillips, J.
JOURNAL    Unpublished.

           DEFINITION Rat salivary proline-rich protein (RP15) gene.
           ACCESSION  M64793 M36414
           VERSION    M64793.1  GI:206711
           SOURCE     Rat (Sprague-Dawley) liver DNA.
           ORGANISM   Rattus norvegicus
                     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
Mammalia;
                     Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
           REFERENCE  1  (bases 1 to 1748)
           AUTHORS    Lin, H.H. and Ann, D.K.
           TITLE      Molecular characterization of rat multigene family
encoding
                     proline-rich proteins
           JOURNAL    Genomics 10, 102-113 (1991)
           MEDLINE    91257817
           FEATURES    Location/Qualifiers
                     source      1..1748
                               /organism="Rattus norvegicus"
                               /strain="Sprague-Dawley"
                               /db_xref="taxon:10116"
                               /tissue_type="liver"
                               /tissue_lib="cosmid genomic library"
                     misc_feature 1802-1810
                               /function=" consensus sequence for initiation in
                               higher eukaryotes "

           FEATURES    Location/Qualifiers
DEFINITION  E. coli periplasmic phosphoanhydride phosphohydrolase (appa)
gene,

           ACCESSION  M58708 L03370 L03371 L03372 L03373 L03374 L03375
           VERSION    M58708.1  GI:145283
           SOURCE     Escherichia coli DNA.
           ORGANISM   Escherichia coli
                     Bacteria; Proteobacteria; gamma subdivision;
Enterobacteriaceae;
                     Escherichia.

           REFERENCE  1  (bases 1811..3109)
           AUTHORS    Dassa, J., Marck, C. and Boquet, P.L.

```

Figure 19A:

```
TITLE      The complete nucleotide sequence of the Escherichia coli
            gene appA reveals significant homology between pH 2.5
            acid phosphatase and glucose-1-phosphatase
JOURNAL    J. Bacteriol. 172 (9), 5497-5500 (1990)
MEDLINE    90368616

FEATURES   Location/Qualifiers
    Source               1811..3109
                        /organism="Escherichia coli"
                        /db_xref="taxon:562"
    sig_peptide          1811..1876
    /gene="appA"
CDS         1811..3109
            /gene="appA"
            /standard_name="acid phosphatase/phytase"
            /transl_table=11
            /product="periplasmic phosphoanhydride
            phosphohydrolase"
            /protein_id="AAA72086.1"
            /db_xref="GI:145285"

    mat_peptide          1877  3106
                        /gene="appA"
                        /product="periplasmic phosphoanhydride
                        phosphohydrolase"

    mutation             replace(1817.. 1819,"gcg changed to gcc")
                        /gene="appA"
                        /standard_name="A3 mutant"
                        /note="created by site directed mutagenesis"
                        /phenotype="silent mutation"
    mutation             replace(3092..3094," ccg changed to ccc")
                        /gene="appA"
                        /standard_name=" P428 mutant"
                        /note="created by site directed mutagenesis"
                        /phenotype=" silent mutation "
    mutation             replace(3095..3097," gcg changed to gct")
                        /gene="appA"
                        /standard_name=" A429 mutant"
                        /note="created by site directed mutagenesis"
                        /phenotype=" silent mutation "
```

Figure 19B:

```
SV40 t intron      3197..3810
                    /note="SV40 signals"
polyA_signal       3807..4047
                    /note="SV40 signals"

BASE COUNT      1257 a      814 c      843 g      1146 t
ORIGIN
   1 GGATCCCCTT TGCTATGTAG TTTTAAATGG AAATTACAAC CCATAGTGTG TTGATAAATA
  61 GAGAGTCCTG TTTGGTTTAA GCAACCTCTG TTTCTCATAA ACTCCATAAA AACAGGAATA
 121 CTCTTTGTTT CTAGCATAAC CAAAAGATTT AGTGAATTGA AAACAATGTT CCCTTAGAGT
 181 ATAGGTCTAA TAACCCCGAA AATATTACCA TGATACTGAG CATTTGTAAG TATCTCATAG
 241 CATGTAGTAT CCATAGTCCA TCAATGAGAG AGACATTTAA CATGATTTTC ATTAATCAGG
 301 TGGAAAAGAC ATGACAACAT TCACAGGCAC TGCACAGAAC ATAGTGGTCC ACCTTGCACA
 361 TATTTCACTA AACTAGGTTT ATCTATTTTG TTGCTTTCTC TAACATCTCT GCAATGAAGC
 421 AGGTCAACAG TGCCACATAT CCTTTACTTA ACCTAAGGAA CACAAAAAAT TTTCTACATA
 481 TATCCTGGTT AGAGAGTGCT TAAAATAAGT TTTCCAAGAA TGGAAAAGAA ATGTTCTGAC
 541 TTAACAATTA AGACAGTATT TATTTAAAGC AAGAAATATG AGGCACACAA GAAAAATATTT
 601 TGGGAAGAAA CCATTTGGTG AACAATATTT CAAATAAAAA TAGACAAACA TAGTTAATTG
 661 TAAAACATAT GTTTGACCAG CCCTTCTTTT CAATAGGCTT AATGTGAATA AAATGTTAAA
 721 GATTCTCTTT GGGTGGCTGC AAATTGTCCA CGAATAAGAC AAAATATAAA AATAAGGACT
 781 GAGTCTCACA AAATGAAAAG GAAATATATT CAGAAAGAGA ATCTTGAGAG AATGTGTTGT
 841 CACAAATTAA AGAAAACCTG TGGTGAATGA CATCCTGAGG CCTGAGCTAT TACTGACATT
 901 TAAGATAAAG GTAACGTAT ACATTTGTCC CATTGAGGGG ACAAGAAAGC TGCTCTCATG
 961 TTCAGCTCTA TAATTCTTGC CTTAAACAAC TTAAATAGAA TGATTTAAAA TATGGAGCTG
1021 TCCATGGACC TTTGAAATAT AAAATAGTCA AGCAACTTAT CAAGGAATTA CAGATTCCTT
1081 GATACTAACA CAGGTAAATC CCACACGTGT TTTGAGACTA CATTGCTGGG GATTTTATTG
1141 ATGTAATAGG TCACATGTTT TTCGGGCCAA TGTGCTGTT ATTGCGTTAC TTCAAGAGAA
1201 TAGTGGCAAC TGATGCTATG TATTCTAGGG GTTTGAAGTG ATGTTTCATG ATTGAAATTT
1261 GTAAAAGAAT AACATCATCA TTCTTAACAA TAGAACATAT AAAGTCACAC AGAAGTGACA
1321 GTGTTTAAGC TGTACTATTG ATCAAAGAAA TTTATTACCT TCAGTTTCAA TGGAAATAAT
1381 TACTGATAAT ACAAACATGT GTGAACACAC ACTAATCCTA TCCAAATGCA CAGTGATACA
1441 CAGAAAATAT TAGCAAGTAG AATGCAATAT TTATATAACG ATTGATTTTA TCAATCAATT
1501 GTATGTATCA ATATATGGGC TATTTTCTTA CACATGATTT TATTCAAATT TACTCTAATC
1561 ATTGTTGAAC CATTTAGAAA AGGCATACTG GCAACTTTTC CTTACCTCAT CCAGCTGGGC
1621 AAAAGTCCCA GTGTGGAGTA AAGGATGCAA GATTTCTGCT TCTGTTAAGT ATAAAAATAAT
1681 AGTATGAATT CAAAGGTGCC ATTCTTCTTG TTCTAGTTAT AAAGGCAGTG CTTGCTTCTT
1741 CCAGACAGA TCTGGATCTC GAGGAGCTTG GCGAGATTTT CAGGAGCTAA GGAAGCTAAA
1801 AGCCGCCACC ATGAAAGCCA TCTTAATCCC ATTTTATCT CTTCTGATT CTTAAACCCC
1861 GCAATCTGCA TTCGCTCAGA GTGAGCCGGA GCTGAAGCTG GAAAGTGTGG TGATTGTCAG
1921 TCGTCATGGT GTGCGTGCTC CAACCAAGGC CACGCAACTG ATGCAGGATG TCACCCGAGA
1981 CGCATGGCCA ACCTGGCCGG TAAACTGGG TTGGCTGACA CCGCGCGGTG GTGAGCTAAT
2041 CGCCTATCTC GGACATTACC AACGCCAGCG TCTGGTAGCC GACGGATTGC TGGCGAAAAA
2101 GGGCTGCCCC CAGTCTGGTC AGGTCGCGAT TATTGCTGAT GTCGACGAGC GTACCCGTAA
2161 AACAGGCGAA GCCTTCGCCG CCGGGCTGGC ACCTGACTGT GCAATAACCG TACATACCCA
2221 GGCAGATACG TCCAGTCCC GATCCGTTATT TAATCCTCTA AAAACTGGCG TTTGCCAACT
2281 GGATAACGCG AACGTGACTG ACGCGATCCT CAGCAGGGCA GGAGGGTCAA TTGCTGACTT
2341 TACCGGGCAT CGGCAAACGG CGTTTCGCGA ACTGGAACGG GTGCTTAATT TTCCGCAATC
2401 AAACCTGTGC CTTAAACGTG AGAAACAGGA CGAAAGCTGT TCATTAACGC AGGCATTACC
2461 ATCGGAACTC AAGGTGAGCG CCGACAATGT CTCATTAACC GGTGCGGTAA GCCTCGCATC
2521 AATGCTGACG GAGATATTTT TCCTGCAACA AGCACAGGGA ATGCCGGAGC CGGGGTGGGG
2581 AAGGATCACC GATTCACACC AGTGGAACAC CTTGCTAAGT TTGCATAACG CGCAATTTTA
2641 TTTGCTACAA CGCACGCCAG AGGTTGCCCG CAGCCGCGCC ACCCGTTAT TAGATTTGAT
2701 CAAGACAGCG TTGACGCCCC ATCCACCGCA AAAACAGGCG TATGGTGTGA CATTACCCAC
2761 TTCAGTCTG TTTATCGCCG GACACGATAC TAATCTGGCA AATCTCGCG GCGCACTGGA
2821 GCTCAACTGG ACGCTTCCCG GTCAGCCGGA TAACACGCCG CCAGGTGGTG AAGGTGTTT
2881 TGAACGCTGG CGTCGGCTAA GCGATAACAG CCAGTGGAAT CAGGTTTCGC TGGTCTTCCA
```

Figure 19C:

```
2941 GACTTTACAG CAGATGCGTG ATAAAACGCC GCTGTCATTA AATACGCCGC CCGGAGAGGT
3001 GAAACTGACC CTGGCAGGAT GTGAAGAGCG AAATGCGCAG GGCATGTGTT CGTTGGCAGG
3061 TTTTACGCAA ATCGTGAATG AAGCACGCAT ACCCGCTTGC AGTTTGTAAG GTATAAGGCA
3121 GTTATTGCTC CCCTTAAACG CCTGGTGCTA CGCCTGAATA AGTGATAATA AGCGGATGAA
3181 TGGCAGAAAT TCGCCGGATC TTTGTGAAGG AACCTTACTT CTGTGGTGTG ACATAATTGG
3241 ACAAACTACC TACAGAGATT TAAAGCTCTA AGGTAAATAT AAAATTTTTTA AGTGTATAAT
3301 GTGTTAAACT ACTGATTCTA ATTGTTTGTG TATTTTAGAT TCCAACCTAT GGAAGTGATG
3361 AATGGGAGCA GTGGTGGAAT GCCTTTAATG AGGAAAACCT GTTTTGCTCA GAAGAAATGC
3421 CATCTAGTGA TGATGAGGCT ACTGCTGACT CTCAACATTC TACTCCTCCA AAAAAGAAGA
3481 GAAAGGTAGA AGACCCCAAG GACTTTCCTT CAGAATTGCT AAGTTTTTTG AGTCATGCTG
3541 TGTTTAGTAA TAGAACTCTT GCTTGCTTTG CTATTTACAC CACAAAGGAA AAAGCTGCAC
3601 TGCTATACAA GAAAATTATG GAAAAATATT CTGTAACCTT TATAAGTAGG CATAACAGTT
3661 ATAATCATAA CATACTGTTT TTTCTTACTC CACACAGGCA TAGAGTGTCT GCTATTAATA
3721 ACTATGCTCA AAAATTGTGT ACCTTTAGCT TTTTAATTTG TAAAGGGGTT AATAAGGAAT
3781 ATTTGATGTA TAGTGCCTTG ACTAGAGATC ATAATCAGCC ATACCACATT TGTAGAGGTT
3841 TTAAGTGCTT TAAAAACCT CCCACACCTC CCCCTGAACC TGAAACATAA AATGAATGCA
3901 ATTGTTGTTG TTAAGTTGTT TATTGCAGCT TATAATGGTT ACAAATAAAG CAATAGCATC
3961 ACAATTTTCA CAAATAAAGC ATTTTTTTCA CTGCATTCTA GTTGTGGTTT GTCCAAACTC
4021 ATCAATGTAT CTTATCATGT CTGGATCGAT CCCCGGGTAC
```

//

Figure 20: Nucleic acid sequence of the known segment of the R15/appa plasmid (including the vector sequences of pBLCAT3 (SEQ ID NO:4).

LOCUS R15/appa 6116 bp DNA SYN 15-APR-2000
 DEFINITION R15/appa transgene with vector
 ACCESSION R15/appa
 REFERENCE 1 (bases 1 to 6116)
 SOURCE synthetic construct.
 ORGANISM synthetic construct
 artificial sequence.
 KEYWORDS salivary proline-rich protein, acid glucose-1-phosphatase; appA
 gene; periplasmic phosphoanhydride phosphohydrolase; artificial
 sequence;
 AUTHORS Golovan, S., Forsberg, C.W., Phillips, J.
 JOURNAL Unpublished.

DEFINITION Rat salivary proline-rich protein (RP15) gene.
 ACCESSION M64793 M36414
 VERSION M64793.1 GI:206711
 SOURCE Rat (Sprague-Dawley) liver DNA.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
 Mammalia;
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 1748)
 AUTHORS Lin, H.H. and Ann, D.K.
 TITLE Molecular characterization of rat multigene family
 encoding
 proline-rich proteins
 JOURNAL Genomics 10, 102-113 (1991)
 MEDLINE 91257817
 FEATURES Location/Qualifiers
 source 1..1748
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /tissue_type="liver"
 /tissue_lib="cosmid genomic library"
 misc_feature 1802-1810
 /function=" consensus sequence for initiation in
 higher eukaryotes "

FEATURES Location/Qualifiers
 DEFINITION E. coli periplasmic phosphoanhydride phosphohydrolase (appA)
 gene,
 ACCESSION M58708 L03370 L03371 L03372 L03373 L03374 L03375
 VERSION M58708.1 GI:145283
 SOURCE Escherichia coli DNA.
 ORGANISM Escherichia coli
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Escherichia.

REFERENCE 1 (bases 1811..3109)
 AUTHORS Dassa, J., Marck, C. and Boquet, P.L.
 TITLE The complete nucleotide sequence of the Escherichia coli gene appA
 reveals significant homology between pH 2.5 acid phosphatase
 and glucose-1-phosphatase
 JOURNAL J. Bacteriol. 172 (9), 5497-5500 (1990)

Figure 20A:

MEDLINE 90368616

FEATURES

	Location/Qualifiers
Source	1811..3109 /organism="Escherichia coli" /db_xref="taxon:562"
sig_peptide /gene="appA"	1811..1876
CDS	1811..3109 /gene="appA" /standard_name="acid phosphatase/phytase" /transl_table=11 /product="periplasmic phosphoanhydride phosphohydrolase" /protein_id="AAA72086.1" /db_xref="GI:145285"
mat_peptide	1877 3106 /gene="appA" /product="periplasmic phosphoanhydride phosphohydrolase"
mutation	replace(1817..1819,"gcg changed to gcc") /gene="appA" /standard_name="A3 mutant" /note="created by site directed mutagenesis" /phenotype="silent mutation"
mutation	replace(3092..3094,"ccg changed to ccc") /gene="appA" /standard_name="P428 mutant" /note="created by site directed mutagenesis" /phenotype="silent mutation "
mutation	replace(3095..3097,"gcg changed to gct") /gene="appA" /standard_name="A429 mutant" /note="created by site directed mutagenesis" /phenotype="silent mutation "

DEFINITION Plasmid pBLCAT3 (bases 3109 to 6116)

ACCESSION X64409

VERSION X64409.1 GI:58163

SOURCE synthetic construct.

ORGANISM synthetic construct
artificial sequence.

REFERENCE 1 (bases 3109 to 6116)

AUTHORS Luckow,B.H.R.

TITLE Direct Submission

JOURNAL Submitted (06-FEB-1992) B.H.R. Luckow, German Cancer Res
Center, Im Neuenheimer Feld 280, W-6900 Heidelberg, FRG

Figure 20B:

REFERENCE 2 (bases 3109 to 6116)
AUTHORS Luckow,B. and Schutz,G.
TITLE CAT constructions with multiple unique restriction sites
for the functional analysis of eukaryotic promoters and
regulatory elements
JOURNAL Nucleic Acids Res. 15 (13), 5490 (1987)
MEDLINE 87260024
COMMENT Promoterless CAT vector for transient transfection
experiments with eukaryotic cells. Allows the analysis of foreign
promoters and enhancers.

FEATURES Location/Qualifiers
source 3109 to 6116
/organism="synthetic construct"
/db_xref="taxon:32630"
polyA_signal 3262..3457
/note="SV40 signals"
CDS complement(4654..5514)
/codon_start=1
/transl_table=11
/gene="Amp"
/product="beta-lactamase"
/protein_id="CAA45753.1"
/db_xref="GI:58165"

BASE COUNT 1724 a 1386 c 1407 g 1599 t
ORIGIN

```
1 GGATCCCCTT TGCTATGTAG TTTTAAATGG AAATTACAAC CCATAGTGTG TTGATAAATA
61 GAGAGTCCTG TTTGGTTTAA GCAACCTCTG TTTCTCATAA ACTCCATAAA AACAGGAATA
121 CTCTTTGTTT CTAGCATAAC CAAAAGATTT AGTGAATTGA AAACAATGTT CCCTTAGAGT
181 ATAGGTCTAA TAACCCCGAA AATATTACCA TGATACTGAG CATTTGTAAG TATCTCATAG
241 CATGTAGTAT CCATAGTCCA TCAATGAGAG AGACATTTAA CATGATTTTC ATTAATCAGG
301 TGGAAAAGAC ATGACAACAT TCACAGGCAC TGCACAGAAC ATAGTGGTCC ACCTTGCACA
361 TATTTCACTA AACTAGGTTT ATCTATTTTG TTGCTTTCTC TAACATCTCT GCAATGAAGC
421 AGGTCAACAG TGCCACATAT CCTTTACTTA ACCTAAGGAA CACAAAAAAT TTTCTACATA
481 TATCCTGGTT AGAGAGTGCT TAAAAAAGT TTTCCAAGAA TGGAAAAGAA ATGTTCTGAC
541 TTAACAATTA AGACAGTATT TATTTAAAGC AAGAAATATG AGGCACACAA GAAAAATATT
601 TGGGAAGAAA CCATTTGGTG AACAATATTT CAAATAAAAA TAGACAAACA TAGTTAATTG
661 TAAAACATAT GTTTGACCAG CCCTTCTTTT CAATAGGCTT AATGTGAATA AAATGTTAAA
721 GATTCTCTTT GGGTGGCTGC AAATGTCCA CGAATAAGAC AAAATATAAA AATAAGGACT
781 GAGTCTCACA AAATGAAAAG GAAATATATT CAGAAAGAGA ATCTTGAGAG AATGTGTTGT
841 CACAAATTAA AGAAAACCTG TGGTGAATGA CATCCTGAGG CCTGAGCTAT TACTGACATT
901 TAAGATAAAG GTAACGTAT ACATTTGTCC CATTGAGGGG ACAAGAAAGC TGCTCTCATG
961 TTCAGCTCTA TAATTCTTGC CTAAACAAC TTAAATAGAA TGATTTAAAA TATGGAGCTG
1021 TCCATGGACC TTTGAAATAT AAAATAGTCA AGCAACTTAT CAAGGAATTA CAGATTCCTT
1081 GATACTAACA CAGGTAAATC CCACACGTGT TTTGAGACTA CATTTGCTGG GATTTTATTG
1141 ATGTAATAGG TCACATGTTT TTCGGGCCAA TGTTGCTGTT ATTCGGTTAC TTCAAGAGAA
1201 TAGTGGCAAC TGATGCTATG TATTCTAGGG GTTTGAAGTG ATGTTTCATG ATTGAAATTT
1261 GTAAAAGAAT AACATCATCA TTCTTAACAA TAGAACATAT AAAGTCACAC AGAAGTGACA
1321 GTGTTTAAAG TGTACTATTG ATCAAAGAAA TTTATTACCT TCAGTTTCAA TGGAAATAAT
1381 TACTGATAAT ACAAACATGT GTGAACACAC ACTAATCCTA TCCAAATGCA CAGTGATACA
1441 CAGAAAATAT TAGCAAGTAG AATGCAATAT TTATATAACG ATTGATTTTA TCAATCAATT
1501 GTATGTATCA ATATATGGGC TATTTCTTA CATATGATTT TATTCAAATT TACTCTAATC
1561 ATGTTGAAC CATTTAGAAA AGGCATACAT GCAACTTTTC CTTACCTCAT CCAGCTGGGC
1621 AAAAGTCCCA GTGTGGAGTA AAGGATGCAA GATTCCTGCT TCTGTTAAGT ATAAAAAAT
```

Figure 20C:

```
1681 AGTATGAATT CAAAGGTGCC ATTCTTCTGC TTCTAGTTAT AAAGGCAGTG CTTGCTTCTT
1741 CCAGCACAGA TCTGGATCTC GAGGAGCTTG GCGAGATTTT CAGGAGCTAA GGAAGCTAAA
1801 AGCCGCCACC ATGAAAGCCA TCTTAATCCC ATTTTATCTT CTTCTGATTC CGTTAACCCC
1861 GCAATCTGCA TTCGCTCAGA GTGAGCCGCA GCTCAAGCTG GAAAGTGTGG TGATTGTCAG
1921 TCGTCATGGT GTGCGTGCTC CAACCAAGGC CACGCAACTG ATGCAGGATG TCACCCAGAG
1981 CGCATGGCCA ACCTGGCCGG TAAACTGGG TTGGCTGACA CCGCGCGGTG GTGAGCTAAT
2041 CGCCTATCTC GGACATTACC AACGCCAGCG TCTGGTAGCC GACGGATTGC TGGCGAAAAA
2101 GGGCTGCCCG CAGTCTGGTC AGGTCGCGAT TATTGCTGAT GTCGACGAGC GTACCCGTAA
2161 AACAGGCGAA GCCTTCGCCG CCGGGCTGGC ACCTGACTGT GCAATAACCG TACATACCCA
2221 GGCAGATACG TCCAGTCCCG ATCCGTTATT TAATCCTCTA AAAACTGGCG TTTGCCAACT
2281 GGATAACGCG AACGTGACTG ACGCGATCCT CAGCAGGGCA GGAGGGTCAA TTGCTGACTT
2341 TACCGGGCAT CGGCAAACGG CGTTTCGCGA ACTGGAACGG GTGCTTAATT TTCCGCAATC
2401 AAAC TTGTGC CTTAAACGTG AGAAACAGGA CGAAAGCTGT TCATTAACGC AGGCATTACC
2461 ATCGGAAC TC AAGGTGAGCG CCGACAATGT CTCATTAACC GGTGCGGTAA GCCTCGCATC
2521 AATGCTGACG GAGATATTTT TCCTGCAACA AGCACAGGGA ATGCCGAGC CGGGGTGGGG
2581 AAGGATCACC GATTACACC AGTGGAACAC CTTGCTAAGT TTGCATAACG CGCAATTTTA
2641 TTTGCTACAA CGCACGCCAG AGGTTGCCCG CAGCCGCGCC ACCCGTTAT TAGATTTGAT
2701 CAAGACAGCG TTGACGCCCC ATCCACCGCA AAAACAGGCG TATGGTGTGA CATTACCCAC
2761 TTCAGTGCTG TTTATCGCCG GACACGATAC TAATCTGGCA AATCTCGGCG GCGCACTGGA
2821 GCTCAACTGG ACGCTTCCCG GTCAGCCGGA TAACACGCCG CCAGGTGGTG AACTGGTGT
2881 TGAACGCTGG CGTCGGCTAA GCGATAACAG CCAGTGGATT CAGGTTTCG TGGTCTTCCA
2941 GACTTTACAG CAGATGCGTG ATAGAACGCC GTGTCTATTA AATACGCCG CGGAGAGGT
3001 GAAACTGACC CTGGCAGGAT GTGAAAGAGCG AAATGCGCAG GGCATGTGTT CGTTGGCAGG
3061 TTTTACGCAA ATCGTGAATG AAGCACGCAT ACCCGCTTGC AGTTTGTAAG GTATAAGGCA
3121 GTTATTGGTG CCCTTAAACG CCTGGTGCTA CGCCTGAATA AGTGATAATA AGCGGATGAA
3181 TGGCAGAAAT TCGCCGGATC TTTGTGAAG AACCTTACTT CTGTGGTGTG ACATAATTGG
3241 ACAAACTACC TACAGAGATT TAAAAACCT CCCACACCTC CCCCTGAACC TGAAACATAA
3301 AATGAATGCA ATTGTTGTTG TTAAC TTGTT TATTGCAGCT TATAATGGTT ACAAAATAAG
3361 CAATAGCATC ACAAAATTTCA CAAATAAAGC ATTTTTTTCA CTGCATTCTA GTTGTGGTTT
3421 GTCCAAACTC ATCAATGTAT CTTATCATGT CTGGATCGAT CCCCCGGTAC CGAGCTCGAA
3481 TTCTTAATCA TGGTCATAGC TGTTTCCTGT GTGAAATTGT TATCCGCTCA CAATCCACA
3541 CAACATACGA GCCGGAAGCA TAAAGTGTA AGCCTGGGGT GCCTAATGAG TGAGCTAACT
3601 CACATTAATT GCGTTGCGCT CACTGCCCCG TTTCCAGTCG GGAAACCTGT CGTGCCAGCT
3661 GCATTAATGA ATCGGCCAAC GCGCGGGGAG AGGCGGTTTG CGTATTGGGC GCTCTTCCGC
3721 TTCCTCGCTC ACTGACTCGC TGCGCTCGGT CGTTCGGCTG CCGCGAGCGG TATCAGCTCA
3781 CTCAAAGGCG GTAATACGGT TATCCACAGA ATCAGGGGAT AACGCAGGAA AGAACATGTG
3841 AGCAAAAGGC CAGCAAAAGG CCAGGAACCG TAAAAAGGCC GCGTTGCTGG CGTTTTTCCA
3901 TAGGCTCCGC CCCCCTGACG AGCATCACAA AAATCGACGC TCAAGTCAGA GGTGGCGAAA
3961 CCCGACAGGA CTATAAAGAT ACCAGGCGTT TCCCCCTGGA AGCTCCCTCG TGCGCTCTCC
4021 TGTTCCGACC CTGCCGCTTA CCGGATACCT GTCCGCCTTT CTCCCTTCGG GAAGCGTGGC
4081 GCTTTCTCAA TGCTCAGCT GTAGGTATCT CAGTTCGGTG TAGGTCGTTT GCTCCAAGCT
4141 GGGCTGTGTG CACGAACCCC CCGTTCAGCC CGACCGCTGC GCCTTATCCG GTAACATCG
4201 TCTTGAGTCC AACCCGGTAA GACACGACTT ATCGCCACTG GCAGCAGCCA CTGGTAACAG
4261 GATTAGCAGA GCGAGGTATG TAGGCGGTGC TACAGAGTTC TTGAAGTGGT GGCCTAACTA
4321 CGGCTACACT AGAAGGACAG TATTTGGTAT CTGCGCTCTG CTGAAGCCAG TTACCTTCGG
4381 AAAAAGAGTT GGTAGCTCTT GATCCGGCAA ACAAACCACC GCTGGTAGCG GTGGTTTTTT
4441 TGTTTGCAAG CAGCAGATTA CGCGCAGAAA AAAAGGATCT CAAGAAGATC CTTTGATCTT
4501 TTCTACGGGG TCTGACGCTC AGTGGAACGA AAACCTACGT TAAGGGATTT TGGTCATGAG
4561 ATTATCAAAA AGGATCTTCA CCTAGATCCT TTAAATTA AAATGAAGTT TTAAATCAAT
4621 CTAAAGTATA TATGAGTAAA CTTGGTCTGA CAGTTACCAA TGCTTAATCA GTGAGGCACC
4681 TATCTCAGCG ATCTGTCTAT TTCGTTTCAT CATAGTTGCC TGACTCCCCG TCGTGTAGAT
4741 AACTACGATA CGGGAGGGCT TACCATCTGG CCCCAGTGCT GCAATGATAC CGCGAGACCC
4801 ACGCTACCG GCTCCAGATT TATCAGCAAT AAACCAGCCA GCCGGAAGGG CCGAGCGCAG
4861 AAGTGGTCCT GCAACTTTAT CCGCTCCAT CCAGTCTATT AATTGTTGCC GGAAGCTAG
4921 AGTAAGTAGT TCGCCAGTTA ATAGTTTGGC AACGTTGTT GCCATTGCTA CAGGCATCGT
4981 GGTGTCACGC TCGTCGTTTGT GTATGGCTTC ATTCAGCTCC GGTTCCCAAC GATCAAGCGC
5041 AGTTACATGA TCCCCCATGT TGTGCAAAAA AGCGGTTAGC TCCTTCGGTC TCCGATCGT
5101 TGTCAGAAGT AAGTTGGCCG CAGTGTTATC ACTCATGGTT ATGGCAGCAC TGCATAATTC
```

Figure 20D:

```
5161 TCTTACTGTC ATGCCATCCG TAAGATGCTT TTCTGTGACT GGTGAGTACT CAACCAAGTC
5221 ATTCTGAGAA TAGTGTATGC GGCGACCGAG TTGCTCTTGC CCGGCGTCAA TACGGGATAA
5281 TACCGCGCCA CATAGCAGAA CTTTAAAAGT GCTCATCATT GGAAAACGTT CTTCGGGGCG
5341 AAAACTCTCA AGGATCTTAC CGCTGTTGAG ATCCAGTTCG ATGTAAACCA CTCCTGCACC
5401 CAACTGATCT TCAGCATCTT TTACTTTTAC CAGCGTTTCT GGGTGAGCAA AAACAGGAAG
5461 GCAAAATGCC GCAAAAAAGG GAATAAGGGC GACACGGAAA TGTTGAATAC TCATACTCTT
5521 CCTTTTTTCAA TATTATTGAA GCATTTATCA GGGTTATTGT CTCATGAGCG GATACATATT
5581 TGAATGTATT TAGAAAAATA AACAAATAGG GGTTCGCGC ACATTTCCCC GAAAAGTGCC
5641 ACCTGACGTC TAAGAAACCA TTATTATCAT GACATTAACC TATAAAAAATA GCGGTATCAC
5701 GAGGCCCTTT CGTCTCGCGC GTTTCGGTGA TGACGGTGAA AACCTCTGAC ACATGCAGCT
5761 CCCGGAGACG GTCACAGCTT GTCTGTAAGC GGATGCCGGG AGCAGACAAG CCCGTCAGGG
5821 CGCGTCAGCG GGTGTTGGCG GGTGTCGGGG CTGGCTTAAC TATGCGGCAT CAGAGCAGAT
5881 TGTA CTGAGA GTGCACCATA TGCGGTGTGA AATACCGCAC AGATGCGTAA GGAGAAAATA
5941 CCGCATCAGG CGCCATTCGC CATT CAGGCT GCGCAACTGT TGGGAAGGGC GATCGGTGCG
6001 GGCCTCTTCG CTATTACGCC AGCTGGCGAA AGGGGGATGT GCTGCAAGGC GATTAAGTTG
6061 GGTAACGCCA GGGTTTTCCC AGTCACGACG TTGTAAAACG ACGGCCAGTG CCAAGC
```

//

Figure 21: Nucleic acid sequence of the known segment of the R15/appa transgene used for the generation of transgenic mice (SEQ ID NO:5).

```

LOCUS      R15/appa      3470 bp      DNA              SYN      15-APR-2000
DEFINITION R15/appa transgene with vector sequences removed.
ACCESSION  R15/appa
REFERENCE  1  (bases 1 to 3470)
SOURCE     synthetic construct.
           ORGANISM  synthetic construct
                     artificial sequence.
KEYWORDS   salivary proline-rich protein, acid glucose-1-phosphatase; appA
           gene; periplasmic phosphoanhydride phosphohydrolase; artificial
           sequence;
AUTHORS    Golovan, S., Forsberg, C.W., Phillips, J.
JOURNAL    Unpublished.

           DEFINITION Rat salivary proline-rich protein (RP15) gene.
           ACCESSION  M64793 M36414
           VERSION    M64793.1  GI:206711
           SOURCE     Rat (Sprague-Dawley) liver DNA.
           ORGANISM   Rattus norvegicus
                     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
Mammalia;
                     Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
           REFERENCE  1  (bases 1 to 1748)
           AUTHORS    Lin, H.H. and Ann, D.K.
           TITLE      Molecular characterization of rat multigene family
encoding
                     proline-rich proteins
           JOURNAL    Genomics 10, 102-113 (1991)
           MEDLINE    91257817
           FEATURES    Location/Qualifiers
                     source      1..1748
                               /organism="Rattus norvegicus"
                               /strain="Sprague-Dawley"
                               /db_xref="taxon:10116"
                               /tissue_type="liver"
                               /tissue_lib="cosmid genomic library"
                     misc_feature 1802-1810
                               /function=" consensus sequence for initiation in
                               higher eukaryotes "

FEATURES             Location/Qualifiers

DEFINITION  E. coli periplasmic phosphoanhydride phosphohydrolase (appA)
gene,

           ACCESSION  M58708 L03370 L03371 L03372 L03373 L03374 L03375
           VERSION    M58708.1  GI:145283
           SOURCE     Escherichia coli DNA.
           ORGANISM   Escherichia coli
                     Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
                     Escherichia.

REFERENCE  1  (bases 1811..3109)
AUTHORS    Dassa, J., Marck, C. and Boquet, P.L.
TITLE      The complete nucleotide sequence of the Escherichia coli gene appA
           reveals significant homology between pH 2.5 acid phosphatase
           and glucose-1-phosphatase

```

Figure 21A:

JOURNAL J. Bacteriol. 172 (9), 5497-5500 (1990)
MEDLINE 90368616

FEATURES

	Location/Qualifiers
Source	1811..3109 /organism="Escherichia coli" /db_xref="taxon:562"
sig_peptide	1811..1876 /gene="appA"
CDS	1811..3109 /gene="appA" /standard_name="acid phosphatase/phytase" /transl_table=11 /product="periplasmic phosphoanhydride phosphohydrolase" /protein_id="AAA72086.1" /db_xref="GI:145285"
mat_peptide	1877 3106 /gene="appA" /product="periplasmic phosphoanhydride phosphohydrolase"
mutation	replace(1817..1819,"gcg changed to gcc") /gene="appA" /standard_name="A3 mutant" /note="created by site directed mutagenesis" /phenotype="silent mutation"
mutation	replace(3092..3094,"ccg changed to ccc") /gene="appA" /standard_name="P428 mutant" /note="created by site directed mutagenesis" /phenotype="silent mutation"
mutation	replace(3095..3097,"gcg changed to gct") /gene="appA" /standard_name="A429 mutant" /note="created by site directed mutagenesis" /phenotype="silent mutation"
polyA_signal	3262..3457 /note="SV40 signals"

BASE COUNT 1065 a 721 c 735 g 949 t

ORIGIN

```
1 GGATCCCCTT TGCTATGTAG TTTTAAATGG AAATTACAAC CCATAGTGTG TTGATAAATA
61 GAGAGTCCTG TTTGGTTTAA GCAACCTCTG TTTCTCATAA ACTCCATAAA AACAGGAATA
121 CTCTTTGTTT CTAGCATAAC CAAAAGATTT AGTGAATTGA AAACAATGTT CCCTTAGAGT
181 ATAGGTCTAA TAACCCCGAA AATATTACCA TGATACTGAG CATTGTGAAG TATCTCATAG
241 CATGTAGTAT CCATAGTCCA TCAATGAGAG AGACATTTAA CATGATTTTC ATTAATCAGG
```

Figure 21B:

```
301 TGGAAAAGAC ATGACAACAT TCACAGGCAC TGCACAGAAC ATAGTGGTCC ACCTTGCACA
361 TATTTCACTA AACTAGGTTT ATCTATTTTG TTGCTTTCTC TAACATCTCT GCAATGAAGC
421 AGGTCAACAG TGCCACATAT CCTTTACTTA ACCTAAGGAA CACAAAAAAT TTTCTACATA
481 TATCCTGGTT AGAGAGTGCT TAAAAATAAGT TTTCCAAGAA TGGAAAAAGAA ATGTTCTGAC
541 TTAACAATTA AGACAGTATT TATTTAAAGC AAGAAATATG AGGCACACAA GAAAAATATTT
601 TGGGAAGAAA CCATTTGGTG AACAATATTT CAAATAAAAA TAGACAAACA TAGTTAATTG
661 TAAAACATAT GTTTGACCAG CCCTTCTTTT CAATAGGCTT AATGTGAATA AAATGTTAAA
721 GATTCTCTTT GGGTGGCTGC AAATTGTCCA CGAATAAGAC AAAATATAAA AATAAGGACT
781 GAGTCTCACA AAATGAAAAG GAAATATATT CAGAAAGAGA ATCTTGAGAG AATGTGTTGT
841 CACAAATTAA AGAAAACCTG TGGTGAATGA CATCCTGAGG CCTGAGCTAT TACTGACATT
901 TAAGATAAAG GTAACGTAT ACATTTGTCC CATTGAGGGG ACAAGAAAGC TGCTCTCATG
961 TTCAGCTCTA TAATTCTTGC CTAAACAAC TTAATAAGAA TGATTTAAAA TATGGAGCTG
1021 TCCATGGACC TTTGAAATAT AAAATAGTCA AGCAACTTAT CAAGGAATTA CAGATTCCTT
1081 GATACTAACA CAGGTAAATC CCACACGTGT TTTGAGACTA CATTTGCTGG GATTTTATTG
1141 ATGTAATAGG TCACATGTTT TTCGGGCCAA TGTTGCTGTT ATTCGGTTAC TTCAAGAGAA
1201 TAGTGGCAAC TGATGCTATG TATTCTAGGG GTTTGAAGTG ATGTTTCATG ATTGAAATTT
1261 GTAAAAGAAT AACATCATCA TTCTTAACAA TAGAACATAT AAAGTCACAC AGAAGTGACA
1321 GTGTTTAAGC TGTACTATTG ATCAAAGAAA TTTATTACCT TCAGTTTCAA TGGAAATAAT
1381 TACTGATAAT ACAAACATGT GTGAACACAC ACTAATCCTA TCCAAATGCA CAGTGATACA
1441 CAGAAAATAT TAGCAAGTAG AATGCAATAT TTATATAACG ATTGTATTTA TCAATCAATT
1501 GTATGTATCA ATATATGGGC TATTTCTTA CATATGATTT TATTCAAATT TACTCTAATC
1561 ATTGTTGAAC CATTTAGAAA AGGCATACTG GCAACTTTTC CTTACCTCAT CCAGCTGGGC
1621 AAAAGTCCCA GTGTGGAGTA AAGGATGCAA GATTTCTCTG TCTGTTAAGT ATAAAAATAAT
1681 AGTATGAATT CAAAGGTGCC ATTCTTCTGC TTCTAGTTAT AAAGGCAGTG CTTGCTTCTT
1741 CCAGCACAGA TCTGGATCTC GAGGAGCTTG GCGAGATTTT CAGGAGCTAA GGAAGCTAAA
1801 AGCCGCCACC ATGAAAGCCA TCTTAATCCC ATTTTTATCT CTTCTGATTG CGTTAACCCC
1861 GCAATCTGCA TTCGCTCAGA GTGAGCCGGA GCTGAAGCTG GAAAGTGTGG TGATTGTCTG
1921 TCGTCATGGT GTGCGTGCTC CAACCAAGGC CACGCAACTG ATGCAGGATG TCACCCGAGA
1981 CGCATGGCCA ACCTGGCCGG TAAAACCTGG TTGGCTGACA CCGCGCGGTG GTGAGCTAAT
2041 CGCCTATCTC GGACATTACC AACGCCAGCG TCTGGTAGCC GACGGATTGC TGGCGAAAAA
2101 GGGCTGCCCG CAGTCTGGTC AGGTCGCGAT TATTGCTGAT GTCGACGAGC GTACCCGTAA
2161 AACAGGCGAA GCCTTCGCCG CCGGGCTGGC ACCTGACTGT GCAATAACCG TACATACCCA
2221 GGCAGATACG TCCAGTCCCG ATCCGTTATT TAATCCTCTA AAAACTGGCG TTTGCCAACT
2281 GGATAACGCG AACGTGACTG ACGCGATCCT CAGCAGGGCA GGAGGGTCAA TTGCTGACTT
2341 TACCGGGCAT CGGCAAACGG CGTTTCGCGA ACTGGAACGG GTGCTTAATT TTCCGCAATC
2401 AAACCTGTGC CTTAAACGTG AGAAACAGGA CGAAAGCTGT TCATTAACGC AGGCATTACC
2461 ATCGGAACCT AAGGTGAGCG CCGACAATGT CTCATTAACC GGTGCGGTAA GCCTCGCATC
2521 AATGCTGACG GAGATATTTT TCCTGCAACA AGCACAGGGA ATGCCGAGC CGGGGTGGGG
2581 AAGGATCACC GATTACACAC AGTGAACAC CTTGCTAAGT TTGCATAACG CGCAATTTTA
2641 TTTGCTACAA CGCACGCCAG AGGTTGCCCG CAGCCGCGCC ACCCGTTAT TAGATTTGAT
2701 CAAGACAGCG TTGACGCCCC ATCCACCGCA AAAACAGGCG TATGGTGTGA CATTACCCAC
2761 TTCAGTGCTG TTTATCGCCG GACACGATAC TAATCTGGCA AATCTCGCG GCGCACTGGA
2821 GCTCAACTGG ACGCTTCCCG GTCAGCCGGA TAACACGCCG CCAGGTGGTG AACTGGTGT
2881 TGAACGCTGG CGTCGGCTAA GCGATAACAG CCAGTGGATT CAGGTTTTCGC TGGTCTTCCA
2941 GACTTTACAG CAGATGCGTG ATAAAACGCC GCTGTCATTA AATACGCCGC CCGGAGAGGT
3001 GAAACTGACC CTGGCAGGAT GTGAAGAGCG AAATGCGCAG GGCATGTGTT CGTTGGCAGG
3061 TTTTACGCAA ATCGTGAATG AAGCACGCAT ACCCGCTTGC AGTTTGTAAG GTATAAGGCA
3121 GTTATTGGTG CCCTTAAACG CCTGGTGCTA CGCCTGAATA AGTGATAATA AGCGGATGAA
3181 TGGCAGAAAT TCGCCGGATC TTTGTGAAGG AACCTTACTT CTGTGGTGTG ACATAATTGG
3241 ACAAACCTACC TACAGAGATT TAAAAAACCT CCCACACCTC CCCCTGAACC TGAAACATAA
3301 AATGAATGCA ATTGTTGTTG TTAACCTGTT TATTGCAGCT TATAATGGTT ACAAATAAAG
3361 CAATAGCATC ACAAATTTCA CAAATAAAGC ATTTTTTTCA CTGCATTCTA GTTGTGGTTT
3421 GTCCAAACTC ATCAATGTAT CTTATCATGT CTGGATCGAT CCCCAGGTAC
```

//

Figure 22: Nucleic acid sequence of the SV40/APPA+intron plasmid (SEQ ID NO:6).

LOCUS SV40/APPA 5421 bp DNA CIRCULAR SYN 14-APR-2000
DEFINITION Ligation of SV40 promoter/enhancer into CAT/APPA+intron
ACCESSION SV40/APPA
REFERENCE 1 (bases 1 to 5421)
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
KEYWORDS SV40 promoter/enhancer, acid glucose-1-phosphatase; appA gene;
periplasmic phosphoanhydride phosphohydrolase; artificial
sequence;
AUTHORS Golovan, S., Forsberg, C.W., Phillips, J.
JOURNAL Unpublished.

DEFINITION E. coli periplasmic phosphoanhydride phosphohydrolase (appA)
gene,

ACCESSION M58708 L03370 L03371 L03372 L03373 L03374 L03375
VERSION M58708.1 GI:145283
SOURCE Escherichia coli DNA.
ORGANISM Escherichia coli
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.

REFERENCE 1 (bases 40 1337)
AUTHORS Dassa, J., Marck, C. and Boquet, P.L.
TITLE The complete nucleotide sequence of the Escherichia coli gene appA
reveals significant homology between pH 2.5 acid phosphatase
and glucose-1-phosphatase
JOURNAL J. Bacteriol. 172 (9), 5497-5500 (1990)
MEDLINE 90368616

FEATURES Location/Qualifiers
Source 40 1337
/organism="Escherichia coli"
/db_xref="taxon:562"
sig_peptide 40..105
/gene="appA"
CDS 40 1337
/gene="appA"
/standard_name="acid phosphatase/phytase"
/transl_table=11
/product="periplasmic phosphoanhydride phosphohydrolase"
/protein_id="AAA72086.1"
/db_xref="GI:145285"
mat_peptide 106 1334
/gene="appA"

Figure 22A:

```

                                /product="periplasmic phosphoanhydride phosphohydrolase"

mutation      replace(46.. 48,"gcg changed to gcc")
               /gene="appA"
               /standard_name="A3 mutant"
               /note="created by site directed mutagenesis"
               /phenotype="silent mutation"
mutation      replace(1320..1322," ccg changed to ccc")
               /gene="appA"
               /standard_name=" P428 mutant"
               /note="created by site directed mutagenesis"
               /phenotype=" silent mutation "
mutation      replace(1323..1325," gcg changed to gct")
               /gene="appA"
               /standard_name=" A429 mutant"
               /note="created by site directed mutagenesis"
               /phenotype=" silent mutation "
```

DEFINITION Plasmid pBLCAT3 (bases 2200 to 4924)

ACCESSION X64409

VERSION X64409.1 GI:58163

SOURCE synthetic construct.

ORGANISM synthetic construct

artificial sequence.

REFERENCE 1 (bases 2200 to 4924)

AUTHORS Luckow,B.H.R.

TITLE Direct Submission

JOURNAL Submitted (06-FEB-1992) B.H.R. Luckow, German Cancer Res

Center, Im Neuenheimer Feld 280, W-6900 Heidelberg, FRG

REFERENCE 2 (bases 2200 to 4924)

AUTHORS Luckow,B. and Schutz,G.

TITLE CAT constructions with multiple unique restriction sites

for

the functional analysis of eukaryotic promoters and

regulatory

elements

JOURNAL Nucleic Acids Res. 15 (13), 5490 (1987)

MEDLINE 87260024

COMMENT Promoterless CAT vector for transient transfection

experiments

with eukaryotic cells. Allows the analysis of foreign

promoters and enhancers.

FEATURES Location/Qualifiers

source 2200 to 4924

/organism="synthetic construct"

/db_xref="taxon:32630"

SV40 t intron 1380..1993

/note="SV40 signals"

polyA_signal 1990..2230

/note="SV40 signals"

CDS complement(3471..4317)

/codon_start=1

/transl_table=11

/gene="Amp"

/product="beta-lactamase"

/protein_id="CAA45753.1"

/db_xref="GI:58165"

Figure 22B:

SV40 promoter/enhancer 5023..5402
/note="SV40 signals"

BASE COUNT 1413 a 1321 c 1331 g 1355 t
ORIGIN

```
1 CGAGATTTTC AGGAGCTAAG GAAGCTAAAA GCCGCCACCA TGAAAGCCAT CTTAATCCCA
61 TTTTATCTC TTCTGATTCC GTTAACCCCG CAATCTGCAT TCGCTCAGAG TGAGCCGGAG
121 CTGAAGCTGG AAAGTGTGGT GATTGTCAGT CGTCATGGTG TCGGTGCTCC AACCAAGGCC
181 ACGCAACTGA TGCAGGATGT CACCCAGAC GCATGGCCAA CCTGGCCGGT AAAACTGGGT
241 TGGCTGACAC CGCGNGGTGG TGAGCTAATC GCCTATCTCG GACATTACCA ACGCCAGCGT
301 CTGGTAGCCG ACGGATTGCT GGCGAAAAAG GGCTGCCCGC AGTCTGGTCA GGTGCGGATT
361 ATTGCTGATG TCGACGAGCG TACCCGTAAA ACAGGCGAAG CCTTCGCCGC CGGGCTGGCA
421 CCTGACTGTG CAATAACCGT ACATACCCAG GCAGATACGT CCAGTCCCGA TCCGTTATTT
481 AATCCTCTAA AAAGTGGCGT TTGCCAACTG GATAACGCGA ACGTGAAGTA CGCGATCCTC
541 AGCAGGGCAG GAGGGTCAAT TGCTGACTTT ACCGGGCATC GGCAAACGGC GTTTCGCGAA
601 CTGGAACGGG TGCTTAATTT TCCGCAATCA AACTTGTGCC TTAAACGTGA GAAACAGGAC
661 GAAAGCTGTT CATTAAACGCA GGCATTACCA TCGGAACCTA AGGTGAGCGC CGACAATGTC
721 TCATTAACCG GTGCGGTAAAG CCTCGCATCA ATGCTGACGG AGATATTTCT CCTGCAACAA
781 GCACAGGGAA TGCCGGAGCC GGGGTGGGGA AGGATCACCG ATTACACCA GTGGAACACC
841 TTGCTAAGTT TGCATAACGC GCAATTTTAT TTGCTACAAC GCACGCCAGA GGTGCCCCG
901 AGCCGCGCCA CCCCCTTATT AGATTTGATC AAGACAGCGT TGACGCCCCA CCACCGCAAA
961 AACAGGCGTA TGGTGTGACA TTACCCACTT CAGTGCTGTT TATCGCCGGA CACGATACTA
1021 ATCTGGCAAA TCTCGGCGGC GCACTGGAGC TCAACTGGAC GCTTCCCCGT CAGCCGGATA
1081 ACACGCCGCC AGGTGGTGAA CTGGTGTTTG AACGCTGGCG TCGGCTAAGC GATAACAGCC
1141 AGTGGATTCA GGTTTCGCTG GTCTTCCAGA CTTTACAGCA GATGCGTGAT AAAACGCCGC
1201 TGTCATTAAA TACGCCGCCC GGAGAGGTGA AACTGACCCT GGCAGGATGT GAAGAGCGAA
1261 ATGCGCAGGG CATGTGTTTCG TTGGCAGGTT TTAGCCAAAT CGTGAATGAA GCACGCATAC
1321 CCGCTTGCGAG TTTGTAAGGC AGTTATTGGT GCCCTTAAAC GCCTGGTGCT ACGCCTGAAT
1381 AAGTGATAAT AAGCGGATGA ATGGCAGAAA TTCGCCGGAT CTTTGTGAAG GAACCTTACT
1441 TCTGTGGTGT GACATAATTG GACAACTAC CTACAGAGAT TTAAAGCTCT AAGGTAAATA
1501 TAAATTTTTT AAGTGATATA TGTGTTAAAC TACTGATTCT AATTGTTTGT GTATTTTAGA
1561 TTCCAACCTA TGGAAGTGAT GAATGGGAGC AGTGGTGGAA TGCCTTTAAT GAGGAAAACC
1621 TGTTTTGCTC AGAAGAAATG CCATCTAGTG ATGATGAGGC TACTGCTGAC TCTCAACATT
1681 CTACTCCTCC AAAAAAGAAG AGAAAGGTAG AAGACCCCAA GGACTTTCCT TCAGAAATTGC
1741 TAAGTTTTTT GAGTCATGCT GTGTTTAGTA ATAGAACTCT TGCTTGCTTT GCTATTTACA
1801 CCACAAAGGA AAAAGCTGCA CTGCTATACA AGAAAATTAT GGAAAAATAT TCTGTAACCT
1861 TTATAAGTAG GCATAACAGT TATAATCATA ACATACTGTT TTTTCTTACT CCACACAGGC
1921 ATAGAGTGTC TGCTATTAAT AACTATGCTC AAAAATTGTG TACCTTTAGC TTTTAAATTT
1981 GTAAAGGGGT TAATAAGGAA TATTTGATGT ATAGTGCCTT GACTAGAGAT CATAATCAGG
2041 CATACCACAT TTGTAGAGGT TTTACTTGCT TTAAAAACC TCCCACACCT CCCCTGAAC
2101 CTGAAACATA AAATGAATGC AATTGTTGTT GTTAACTTGT TTATTGCAGC TTATAATGGT
2161 TACAAATAAA GCAATAGCAT CACAAATTTT ACAAATAAAG CATTTTTTTC ACTGCATTCT
2221 AGTTGTGGTT TGTCCAAACT CATCAATGTA TCTTATCATG TCTGGATCGA TCCCCGGGTA
2281 CCGAGCTCGA ATTCGTAATC ATGGTCATAG CTGTTTCCTG TGTGAAATTG TTATCCGCTC
2341 ACAATTCAC ACAACATACG AGCCGGAAGC ATAAAGTGTA AAGCCTGGGG TGCCTAATGA
2401 GTGAGCTAAC TCACATTAAT TGCGTTGCGC TCACTGCCCG CTTTCCAGTC GGGAAACCTG
2461 TCGTGCCAGC TGCATTAATG AATCGGCCAA CGCGCGGGGA GAGGCGGTTT GCGTATTGGG
2521 CGCTCTTCCG CTTCTCGCT CACTGACTCG CTGCGCTCGG TCGTTCGGCT GCGGCGAGCG
2581 GTATCAGCTC ACTCAAAGGC GGTAATACGG TTATCCACAG AATCAGGGGA TAACGCAGGA
2641 AAGAACATGT GAGCAAAAGG CCAGCAAAAG GCCAGGAACC GTAAAAAGGC CGCGTTGCTG
2701 GCGTTTTTCC ATAGGCTCCG CCCCCCTGAC GAGCATCACA AAAATCGACG CTCAAGTCAG
2761 AGGTGGCGAA ACCCGACAGG ACTATAAAGA TACCAGGCGT TTCCCCCTGG AAGCTCCCTC
2821 TGTGCGCTCT CTGTTCCGAC CTGCGCGCTT ACCGGATACC TGTCCGCTT TCTCCCTTCG
2881 GGAAGCGTGG CGCTTTCTCA ATGTCTACGC TGTAGGTATC TCAGTTCGGT GTAGTTCGTT
2941 CGCTCCAAGC TGGGCTGTGT GCACGAACCC CCCGTTTCAGC CCGACCGCTG CGCCTTATCC
3001 GGTAACATAT GTCTTGAGTC CAACCCGGTA AGACACGACT TATCGCCACT GGCAGCAGCC
3061 ACTGGTAACA GGATTAGCAG AGCGAGGTAT GTAGGCGGTG CTACAGAGTT CTTGAAGTGG
```

Figure 22C:

```
3121 TGGCCTAACT ACGGCTACAC TAGAAGGACA GTATTTGGTA TCTGCGCTCT GCTGAAGCCA
3181 GTTACCTTCG GAAAAAGAGT TGGTAGCTCT TGATCCGGCA AACAAACCAC CGCTGGTAGC
3241 GGTGGTTTTT TTGTTTGCAA GCAGCAGATT ACGCGCAGAA AAAAAGGATC TCAAGAAGAT
3301 CCTTTGATCT TTTCTACGGG GTCTGACGCT CAGTCCAACC AAAACTCACG TTAAGGGATT
3361 TTGGTCATGA GATTATCAAA AAGGATCTTC ACCTAGATCC TTTTAAATTA AAAATGAAGT
3421 TTTAAATCAA TCTAAAGTAT ATATGAGTAA ACTTGGTCTG ACAGTTACCA ATGCTTAATC
3481 AGTGAGGCAC CTATCTCAGC GATCTGTCTA TTTTCGTTTCAT CCATAGTTGC CTGACTCCCC
3541 GTCGTGTAGA TAACTACGAT ACGGGAGGGC TTACCATCTG GCCCCAGTGC TGCAATGATA
3601 CCGCGAGACC CACGCTCACC GGCTCCAGAT TTATCAGCAA TAAACCAGCC AGCCGGAAGG
3661 GCCGAGCGCA GAAGTGGTCC TGCAACTTTA TCCGCCTCCA TCCAGTCTAT TAATTGTTGC
3721 CGGGAAGCTA GAGTAAGTAG TTCGCCAGTT AATAGTTTGC GCAACGTTGT TGCCATTGCT
3781 ACAGGCATCG TGGTGTACAG CTCGTCGTTT GGTATGGCTT CATTAGCTC CGGTTCCTAA
3841 CGATCAAGGC GAGTTACATG ATCCCCCATG TTGTGCAAAA AAGCGGTTAG CTCCTTCGGT
3901 CCTCCGATCG TTGTCAGAAG TAAGTTGGCC GCAGTGTTAT CACTCATGGT TATGGCAGCA
3961 CTGCATAATT CTCTTACTGT CATGCCATCC GTAAGATGCT TTTCTGTGAC TGGTGAGTAC
4021 TCAACCAAGT CATTCTGAGA ATAGTGTATG CGGCGACCGA GTTGCTCTTG CCCGGCGTCA
4081 ATACGGGATA ATACCGCGCC ACATAGCAGA ACTTTAAAAG TGCTCATCAT TGGAAAACGT
4141 TCTTCGGGGC GAAAACCTCTC AAGGATCTTA CCGCTGTTGA GATCCAGTTC GATGTAACCC
4201 ACTCGTGCAC CCAACTGATC TTCAGCATCT TTTACTTTCA CCAGCGTTTC TGGGTGAGCA
4261 AAAACAGGAA GGCAAAATGC CGCAAAAAG GGAATAAGGG CGACACGGAA ATGTTGAATA
4321 CTCATACTCT TCCTTTTTCA ATATTATTGA AGCATTTATC AGGGTTATTG TCTCATGAGC
4381 GGATACATAT TTGAATGTAT TTAGAAAAAT AAACAAATAG GGGTTCCGCG CACATTTCCC
4441 CGAAAAGTGC CACCTGACGT CTAAGAAACC ATTATTATCA TGACATTAAC CTATAAAAAT
4501 AGGCGTATCA CGAGGCCCTT TCGTCTCGCG CGTTTCGGTG ATGACGGTGA AAACCTCTGA
4561 CACATGCAGC TCCCGGAGAC GGTACAGCT TGTCTGTAAG CGGATGCCGG GAGCAGACAA
4621 GCCCCTCAGG GCGCGTCAGC GGGTGTGGC GGGTGTGGG GCTGGCTTAA CTATGCGGCA
4681 TCAGAGCAGA TTGTACTGAG AGTGCACCAT ATGCGGTGTG AAATACCGCA CAGATGCGTA
4741 AGGAGAAAAT ACCGCATCAG GCGCCATTCG CCATTCAGGC TGCGCAACTG TTGGGAAGGG
4801 CGATCGGTGC GGGCCTCTTC GCTATTACGC CAGCTGGCGA AAGGGGGATG TGCTGCAAGG
4861 CGATTAAGTT GGGTAACGCC AGGGTTTTCC CAGTCACGAC GTTGTAATAAC GACGGCCAGT
4921 GCCAAGCTTT ACACCTTATG CTTCCGGCTC GTATGTTGTG TGGAATTGTG AGCGGATAAC
4981 AATTTACAC AGGAAACAGC TATGACCATG ATTACGAATT CGGCGCAGCA CCATGGCCTG
5041 AAATAACCTC TGAAAGAGGA ACTTGTTTAG GTACCTTCTG AGGCGGAAAG AACCAGCTGT
5101 GGAATGTGTG TCAGTTAGGG TGTGGAAGT CCCCAGGCTC CCCAGCAGGC AGAAGTATGC
5161 AAAGCATGCA TCTCAATTAG TCAGCAACCA GGTGTGGAAG GTCCCCAGGC TCCCCAGCAG
5221 GCAGAAGTAT GCAAAGCATG CATCTCAATT AGTCAGCAAC CATAGTCCCG CCCCTAACTC
5281 CGCCCATCCC GCCCTAACT CCGCCCAGTT CCGCCCATT CCGCCCCAT GGCTGACTAA
5341 TTTTTTTTAT TTATGCAGAG CCCGAGGCCG CCTCGGCCTC TGAGCTATTC CAGAAGTAGT
5401 GAGGAGGCTC GAGGAGCTTG G
```

//

Figure 23. The nucleic acid sequence of the Lama2/APPA transgene used for the generation of transgenic mice and transgenic pigs (SEQ ID NO: 7)

LOCUS transgene 17732 bp DNA SYN 14-APR-2000
 DEFINITION Lama-appA cut XhoI..20623 to NotI..17732
 ACCESSION transgene
 KEYWORDS parotid secretory protein; acid glucose-1-phosphatase; appA gene;
 periplasmic phosphoanhydride phosphohydrolase; artificial sequence;
 cloning vector
 REFERENCE 1 (bases 1 to 17732)
 AUTHORS Golovan, S., Forsberg, C.W., Phillips, J.
 JOURNAL Unpublished.

FEATURES

DEFINITION M. musculus Psp gene for parotid secretory protein.
 ACCESSION X68699
 VERSION X68699.1 GI:53809
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 3777 to 5332;)
 AUTHORS Svendsen, P., Laursen, J., Krogh-Pedersen, H. and Hjorth, J.P.
 TITLE Novel salivary gland specific binding elements located in
 the PSP proximal enhancer core
 JOURNAL Nucleic Acids Res. 26 (11), 2761-2770 (1998)
 MEDLINE 98256451
 REFERENCE 2 (bases 7147 to 12653; 13952 to 17731)
 AUTHORS Mikkelsen, T.R.
 TITLE Direct Submission
 JOURNAL Submitted (07-OCT-1992) T.R. Mikkelsen, Department of
 Molecular Biology, University of Aarhus, CF Mollers Alle
 130, 8000 Aarhus, DENMARK
 REFERENCE 3 (bases 7147 to 12653; 13952 to 17731)
 AUTHORS Laursen J, Hjorth JP
 TITLE A cassette for high-level expression in the mouse salivary
 glands.
 JOURNAL Gene 1997 Oct 1;198(1-2):367-72
 MEDLINE 9370303

FEATURES

Location/Qualifiers

source 1.to 12653; 13952 to 17731
 /organism="Mus musculus"
 /strain="C3H/As"
 /db_xref="taxon:10090"
 /chromosome="2"
 /map="Estimate: 69 cM from centromere"
 /clone="Lambda YP1, Lambda YP3, Lambda YP7"
 /clone_lib="Lambda-PHAGE (Lambda L47.1)"
 /germline
 /note="Allele: b"

misc_feature 3777-5332
 /gene="PSP"
 /function="salivary gland specific positive acting
 regulatory region"
 enhancer 7147..8724

Figure 23A:

```

      /evidence=experimental
exon    11778..11824
      /gene="Psp"
      /note="exon a"
      /number=1
      /evidence=experimental
exon    12626.. 14190
      /gene="Psp"
      /note="exon b fused with exons h and i"
misc_feature    12644-12652
      /function=" consensus sequence for initiation in higher
      eukaryotes ."
misc_feature    13952-13965
      /function=" M13mp18 polylinker"
```

DEFINITION E. coli periplasmic phosphoanhydride phosphohydrolase (appA) gene,

```

ACCESSION    M58708 L03370 L03371 L03372 L03373 L03374 L03375
VERSION      M58708.1  GI:145283
SOURCE       Escherichia coli DNA.
ORGANISM      Escherichia coli
              Bacteria; Proteobacteria; gamma subdivision;
              Enterobacteriaceae;
              Escherichia.
```

```

REFERENCE    1  (bases 12653..13951)
AUTHORS      Dassa,J., Marck,C. and Boquet,P.L.
TITLE        The complete nucleotide sequence of the Escherichia coli
              gene appA reveals significant homology between pH 2.5
              acid phosphatase and glucose-1-phosphatase
JOURNAL      J. Bacteriol. 172 (9), 5497-5500 (1990)
MEDLINE      90368616
```

```

FEATURES      Location/Qualifiers
Source          12653..13951
                /organism="Escherichia coli"
                /db_xref="taxon:562"
sig_peptide     12653..12718
/gene="appA"
CDS             12653      13951
                /gene="appA"
                /standard_name="acid phosphatase/phytase"
                /transl_table=11
                /product="periplasmic phosphoanhydride
                phosphohydrolase"
                /protein_id="AAA72086.1"
                /db_xref="GI:145285"
```

Figure 23B:

```
mat_peptide      12719 13948
                  /gene="appA"
                  /product="periplasmic phosphoanhydride
                  phosphohydrolase"

mutation  replace(12659.. 12661,"gcg changed to gcc")
           /gene="appA"
           /standard_name="A3 mutant"
           /note="created by site directed mutagenesis"
           /citation={3}
           /phenotype="silent mutation"

mutation  replace(13934..13936," ccg changed to ccc")
           /gene="appA"
           /standard_name=" P428 mutant"
           /note="created by site directed mutagenesis"
           /citation={3}
           /phenotype=" silent mutation "

mutation  replace(13937..13939," gcg changed to gct")
           /gene="appA"
           /standard_name=" A429 mutant"
           /note="created by site directed mutagenesis"
           /citation={3}
           /phenotype=" silent mutation "
```

```
BASE COUNT      4719 a   4125 c   4168 g   4719 t
ORIGIN
   1 TCGAGAGTAT CTTTGTTCAGC TGTGCCTCCA ACAAAGGGGT ACTGTTGCCC ACATAGAAAG
  61 ATCTAAACTA ATTAATTAAT CCCTCACCCG CAAATCTTTC AGTCACTAAG TTAGCACGAT
 121 TGTGTAACAA GTTCTCCAAA GGAGAGATAC AGATGAGTGC GTATAGGGTG GACCTGGCTG
 181 CTGAGGAGAC ACCTGCATCT GACTAAGAAG AGCCACGGTG TTAGTTGAAT GGTGTGGAGT
 241 AGGGTGGTTC TGTGGGACAG TAGAAAATCG AGAGGCATGT GCCGTTTAGT GAACTGATGG
 301 AAGCTACCCC AAACGACAGA GATTGTTCAGT CAGGCCAATC CGTTTCGAGT TTGATGGGCA
 361 GCCGGACAGT GAGACAGACA CACCTACTCA GTTGGAGGAA GGATGAGAAC AATGGCCAGC
 421 AGGGATTGAG AGACCCTGAC AGGCGCAAGG CCCTAACACA CACACCTACC ACCTCACTTG
 481 ACAAAGCTGC CAAAGACCAA AGACTTGTTT TCATTAGAA ATGACAGCTG GCTTGACCCG
 541 ACAGCATAAT AAGCAGAGTG TACTCTGATT GGAGAACTTT AATGTGTTTC ATTCACTATT
 601 ATAAAAGGAC AGTATTACAG ATTTTGTTGT ACACTGCTGT TACATGTGGG GCAGTGTGTC
 661 TTTAAGTAGG GTAAAGTACT CTTTAAAAAT GGGTCCTAGA TATTTTTTCC TTTAACTCAA
 721 GTCTCTTACT GTTTAAATGA TTTTATTTT GTTTAATATG GAGGAAAAAG AAGCGTAAAT
 781 GGACAATATA TATTTAGAGA AAGATGGTTA GCTGTCAGAA AAATATGCAA ATCAAAATCA
 841 CACCAAGACT GCAGCACACC CTTGTCAGAT GGCTGTGATC AAGAAAATAA ATGACAATGA
 901 GTGGTGGTGA AGATGTACTA AAGGGAAACA CACACACACA CACACACACA CACACACACA
 961 CCACTGGAG CAACCACTGT GGAAATCAGT ATGAATGGTC CTCAAAAACC TGAAGATAGA
1021 GCGGGGCGTG GTGGCATACA CTTTATTTCC CAGCACTGGG GAGGCAGAGG CAGGTGGATC
1081 TCTGAGTTCC AGGCCAGCCT GGTCTATAGC ACAGGTTCTA GGACAGCCAG GGCTACACAG
1141 AAAAACCCCTG CCTTGATTAA ACCAAACCAA ACCAAACCAA ACCAAACCAA ACCAAACCAA
1201 ACCAAACCAA ACCAAACCAG ACCAAACCAA AACACTGAAG ATAGAACTTC AGTATTCCAT
1261 TCCTAGATAT ATACCCAATG GAGACTAAGT CAGCAAGACA CCTGCACAGC CATGTTCACT
1321 ACTACACTGT TCACCACAGC CAGGCTGTGG AACCAGCCTG AGTGTCCATG ATAAATGAAT
```

Figure 23C:

1381 GGATAGGTAA CTTTCAAGGT AAATGGACTC TGCTGTGTAC ATGCCTCACA TTCFGTTTAT
1441 TCATTTTTCT TTATGAGGTG TCCATTGAGG AGTCACATGG TAGTTCTATT TTCAGTCTTC
1501 TGAAGATACT ACACTGGTCC CCACAGTTTA CACTTTTATC AGCAGTGAAT AAGGGTTCCT
1561 CTATCCTTAC CATCATTTGT TGTAATTTTT CTTGATGACC CTCTTTCTGA CAGGGATAGG
1621 ATGTAATATC AGTGTGAGGA AGTACAACCT GTTTTCTAAG TATTTATTGG CCCCTTGCAT
1681 TTCTTCTTTT GAAAACCTGTC GGTCCTGAC ATCTGCTCAG GTATTCATTG GATGTTGTTT
1741 CTTTGGTGTT TGAGTTCTTA TGAATTCTAG ATGTTAAATC CCTGCCTGTG GTTCTCTCCC
1801 ATTCTGTAGG CTGCCTCCTC ACCCTGGCAA TTGTTGTCCT TGTTTTCAG AAACTTTTGA
1861 CTTTCATGGA TCTCATTTGT CAGTTTTCCC TCCTCTGCTA TAGCCTGAGC TAATGCACTG
1921 GTTTTTACAG AGCCCTGGTC TAGCCCTTTA TCCTCCTCTG GCAGCTTCGG AGTTTCATTT
1981 CTTACATTTA GATCTTTGAT CCACTTTGAA CAAGTTTGG AGCAGGGTGA GAGATACGAA
2041 TCTAGTTCCA TTCTTCCATA TGTGATCCTA GTTTACATAG CATCGTTGGT TGAAGAGGTT
2101 TTATTTTTATT TTAAATAAT GTGTCATAAA AAACGAGGTG GTTGTAGCAG TGTGGATTG
2161 TTTCTTTGTC CTTTGATCTA CAGGTCCTGT TTTGTGTCAG TCTCATGATG TTTTATTGCT
2221 ATGGCTCTGT CATAAGTCT GAGGTCAGGT ATTGTGATAT ACCTTCAGTA TTGCTCCCTC
2281 AGACTCAGGT TTGCTTTGGC CAGGAGTCAT CTACTCAGT GCTCTTAGAG CTCCCCCAGC
2341 ATGTAGCTGC TACTATTCTT AGTTGATAAA TCAGGAAACT GGGGCTCAGA GAGATTAACT
2401 GTCTTGAAC TCTTCTGGGG AGGTGAAACG TGGAGACACT AAACGTGTGT TACCCTGTAC
2461 TGCTCCAGTA GCTGTGCGGT GCTGGGTCTAC AGCAAAGCAC CTATACTATA TATTACTCAG
2521 GAGGTGGAAG AACTCAGCCT CCCTTGGGGT TCCCAAGCTC CCAGGTGTCC AGTCACTGCT
2581 GGAAACCTCA TGGAGTCTGA AAGGAAGGTT TGAGGGTACA TGGGGCAGCG ATGAGGAGCC
2641 TGGGGCTGGG ATCTCCCAAA CACCTGGATA TCCAGATGCC ACTGGGTGAG GGGGAGTTGG
2701 GAACAGAGTT GGGATGTCCA TGGACCTGTG ACAAGGCCAG GGCCAGGGGG AGGATAACTC
2761 TGGCTTTACT AATTTGCGAA AGTCCTTAGC TTAGCAGCAG TTGTCTGGGA GCACAGAGGG
2821 GCCTTCTGTA AGAGGCTCAG GCAGTGCCGC TCTGTAGGCG AAGGTCTTCT CCATGTTCCC
2881 CATGGTGGTT CTTGATGAAA GAGACAGTCC TTGGCTCCAA ACTGGTTTAT TGATTGTTCA
2941 TTGTGGAAAA TGGGTGCACA CCACCTTCTC AGGGTGGACC AGAGATCAAA TACCTTTTGC
3001 AGGGAGGAAT ATCTGGGAAG GGACGCTTAC TGGCTAAACC CTCAGGGCCT CTAGATACAT
3061 CATTAGCATG GAGAACTCTG TTCTGGGCTA CATGACCACA GGCCACATTT CCACAAGCCA
3121 CATGTGGGAA GTGTGGCACA TGTCTAGGC CAGGAATCTG GTAGGGAGCG TGGAGCCACC
3181 TACCATCCCA GGTGGGTGCC TGGGTGCCAG GGACCCTGAA CCCGCTCAAC CTTACCAAGT
3241 TTCCTGGCAG GGTCCACTGT CCTACACAGA AGCTGGAGGA GGTGTGAGGG TTGTGTCTTT
3301 GTGGAATGTC CCATGCTGCT TGGGGCTCAG TTTCTCCACC TGTACCTCAT TGGTTTGGGT
3361 ATAAAAAGTG GGGATACTTT ATTATTCTCT GACTCGGTCC TGAGGAAAAA GCATCGTGGC
3421 AGTCCAGGAA CCACACCCTG AGGTTCCCTG ACTGAAGGGA CTCCCTAAGT CTCTGGAGTC
3481 TCTCCCCTTC ACAGAGCTGC CAAAGTCTAG GTTCTTTTGA GGATAACAGA GCCATGCTTG
3541 GTAAGCAGAC AACAGCATTT GTTTACTCAA CCTTCTTTTG TCAGCTCCCT GTTCATAAAC
3601 AAGTTGAGAC ACCATGCTGG CTTGAGGAAG ACTTCTAAAG CCAGACAACCT GTGCAAGGAA
3661 GAAGAAGAAG GGGCAAGTGG AGTTAGCCTG GATGTAGCCC TCAAAGTCTC CAGAGACCAG
3721 CCATGAAGGC TCAAGTGGAG GGCAAGACCT GCAGCAGCCA AGCATCTGGC AGGAGAGGAT
3781 CCTGGGAACC CCTCTACCAT GACACACATT CTTCTGTCAG GTCACACTTA ATAGGCCATT
3841 TCTTATTTGG ATCTATCATG GTGTTCTGTG CGAGATTAAT GAGGTGTTAT GCTGCGAACA
3901 GAAAGTTATA TAAAAACAAG TCCCCCCCCC TTGTCACTGC TGCTAAGAAT GTAGCAGAAA
3961 TTGTCTCAAG TGTCTCTCTA ATCAGAAACA ATAAAGGTCT CCTTGGATTG AAGCCCTCCA
4021 GTTTCCTCCT TCCTTGCTGA GCCTTGGACA CCCATACAAA CCTCCTGGAT GCTACAGCTC
4081 TGGGCAGAGA CTCCAAGGTG GGGAGAGACT GATGGTACAA AAGCAAAATA CTGTGTTGGG
4141 GGTACACCCA CTCCTCTGCC TGTGTGGTTC CTGCAGTCAG TCCTGCAGAC AGGCCCTCAG
4201 TGGGTCTTCC ATGGGCAACA CGCAGAGGGA GGCAATGGAT GGAATACCC ACACCCTGGT
4261 TAGTTTACCC CGGCCATGCT CTCTGCTCTT CATCCCTCCT CTGCCCTCTG CCACGGCTTT
4321 CTCTGCAGGA ATCATATCTT CATATTGGCC CACAGGTGTT CTCCTCAGCC TAGCTATGAT
4381 GTTTACTTTA GAGTGACCTT AGCAGGGCTG GTGGGAATGA GTTCTAGAAG GCTCAGGGAG
4441 ATGCTAGGGA AGAAACGTCT TCTAACTACT GAGGTTACTA AGTTCCTGGT GGTGTCTCT
4501 GCCTTTCCCT TGTTAAAGTC ACCTTGAAGT TAGTGCAGAA GAAATCAGAG CCCAGTCACA
4561 GAGTAAATAT GGTCTGAAG ATTTCTTTTG AGTGCCCGA ATCCATGACA TTTCAAGAGC
4621 CCTCTTGTGA CCTTAAGTCA TTTGGGGTTG TATCTTCTGC TTGATGTATG TGTGTGTGTT
4681 TATCAAAGAG TGAGATGGTT ACATAAGAGG TGCTCTAAAG GACAGAGAGG ATTTGCAATT
4741 GTGGCATGTG ACATCCTCAG GCCTTGCTCT GGTGCCAGGA GGAAGTGTG CAGAAAAGAG
4801 TAAGAGGTCA TTTCTGGAG GCTGTCACTA TAGAGGAGAT CTTAGAGTGC ATTCCCTCCT

Figure 23D:

4861 CCAGGCCCTG CCTGAGGATA GACATGTGCT GACTGCAACT GAAACAGAGG CTTGGGATGG
4921 AGAGTTAGGT TCACAGAAGG GAGGGTGGGA GATGGATGCT TGCTGGGTTTC TGGGTCTCAT
4981 CACCAGCTCC TGACCACCCG GTCAGCCCAT GTGCTTATTC CATAGCTTTC TTTTGCTATG
5041 TTTACTCAGT GTGGTGTITG TTGGGACCCA GCAGAAGCCA GTCCCAGGCT GACAGCTGTG
5101 GATACACAGG GCAGCATGAG GGTCCCTCAGC CTGAAGCAGT CAGGCTGGCA GAAGAGAAAG
5161 ACCAGCACAC ATTCCCTCAA CCAACTATGT CTTGAAAAAC AAACATATTA TATCACATAT
5221 ATTGCATTTA TGAGACAGCT AAAATGTACT CGGGTAGCAT GACTCCAGGT GGGGATATCT
5281 GCAAGTGCCA TGAGTGGCAG AGGGACAGCC AATGTGAGGC AAGAAGGAAT TCTGGCTCAA
5341 CACAGCTTAG CTCCCTGGTG TTGGTTCAAA CTTTGAGAGT TTGACCACAA GCACTTTATT
5401 TTTGACATAT TTAAACAGAG CACAACCTTG GGAAAAAGTT TTCTTATGAA AATTATCACA
5461 ATAAAGCTTA AGGCATGACT ACATTAATAAT GCCTTTGCAA AGTATATGTG CCCTCTTCCA
5521 CAAGAATGGT TCTATTGACT GAGAAATAAT GTTCAGGATA AAGATCCAGG AAGAAAAGAT
5581 CAGGGATAAG TAAAATACTA AACTCTTTTG CAAAGTACAT AGACCCTCTT TCATAACAAT
5641 GGGTTCTATT GACTGACAAG CACTGCTCAG GAGTTGGGAA AGAGTCTAGC ATAAGCACGA
5701 TAGCCTGGAG ACTCTAGTGA GGTCTAGTCT TACAGACAGC AAAAATCACC AGGTTACAAA
5761 CTACATTCAT TTCCAGTTTT CTGATCAGGC ACAGGTATGA ATCCCTTCTG TTGAAGAGAA
5821 AAGTCCATGT GTTTAAATA TCTGGTTTCT CCAGTGCTAT TAGCGAGAAG ACTTGAGCCC
5881 TATACAACCT CCACCTGGAG TGACATCCTG TCTTCATGGT ATATTACATA CTTAGACACG
5941 CTCATCTCAG AGACTTAGGA CTTTGTCTTC TGATCTCCAT TTCTGATCCC ACTTCCACCT
6001 TTGCCTTGAT AGTGTCAATT TCTTCACTGC CTTGGTGACA ACCATGTTAT CCTCTGTGTA
6061 TTTGAGTGTT ACCATTTTCA GATTTTACCT GTATGCAAGA TCACACAGTC TTTGTCTTTC
6121 TGTCTGGATG CATGCTAATC TCTACACAAC AACCCTTCCC CGTCACTCAG ATCTTCTTCC
6181 ATTAACACAT ACATGGTGCT GAAGAGGCTA GGGAGCTTCC CTTCACTGGG GAGCTAGCTG
6241 GCTATTGGGC CTTTTTGACT GTCCAGGAAG GCCCCCAATT GCTGAGACAA GAACTTAGAT
6301 TCTTCATTAT TGACTCTAAC TCATGTATCA AGCAGAAGCT AATGAATAGT TATCAACAGG
6361 ATCAGAGGTT CCAGTGTAAG ACACCTTGAC ATGAAAGAAC GGAGGAAGGA CAGATGGATG
6421 CATAAAAGCA GGACCACTGC CCCAGGAAGG TCCTGGAAAC TGATGCAGGG CAAAGGCACG
6481 GTTATAAACC AAATCTTAGG GAGTCAGGAA GAGCACAGAG GAGCTCAACC AACTGACCAC
6541 TGCCTTAGGGG GTACCAACCC AATCCTCCCT GTGGGAACAG CTAAGCTATC AGCCAAGGGT
6601 AATAAACAGG CAGGACCTGT GGATGACATG GAGAGCATAG GGACCCTGGG TCCAGCCTTT
6661 AGCACCTGCA CTCTCAGGAT ACTCCACCAT TGTGTCTTAG AGAGCCTAGG GATACTGGGT
6721 CCAGCCTTTG GTACCTTCAC TCTCAGGTA CCCCATCACT GTGTCTTGA GAGCCTAGGC
6781 ACCCTGGGTC CAGCCTTCAG TACCTGCGCT CTCAGGACAC CCCACCATG TCTCTTGCCC
6841 CGTCTCTTCT TCCTCTTCCT CCCTTTTCACT GTCTCTTCTC TGTTTCTTTC TTGACTCTCC
6901 TTTCCCTCA CACCCTCACT CTAGTTCTCC CTTTCCCTCT CTGCATCACC CTATTCTCTC
6961 TGTGGTCCCT CCACTTTCCT TTATCTCTCA TGCTTCTCTC CTCCCTCAA TCTTGTACAC
7021 CCACATATACT TCAGGGGCCA GCTCTAGTGA CAAAGCTGTT AATAGCAAGA CTCTCAGATC
7081 TCCAACGCT CAGAGGAGCC AGACCCACCA AGAAGCTCTT CCAGGTCCAA TTTCAAGTTC
7141 CTTGAAAGC TTTCAGCAAA TGCTCAGGGA ACATGCCACT AACAAGAAGA TGCAAATFCC
7201 AGTTGAGAGT GGGAAAGGCC CTTGCGTAGG TCCATCTTC CAGGCCAAGG TCAGAGGGGC
7261 TCTGTGTAAT CCGGATTGAC AGGGCTCAGA ACAATGTTTT GTTTTAAAGG TTTATTTATT
7321 TTAGGTGTTA GTGTCTTTGC TTGCATGACC TTATGTGCAT CATGTGTGTG CAGGTTCTCTG
7381 ATGACAGTAG AGGAGGGCTT TGAATCCCTG GGGATAGGAA GTTACAGGAA ATTATAAGCT
7441 GCTTTGTGGG TCTTCTAGCT TTCCCAACAG AAGTGAATGC TCTTACCAC TGAGCCATCT
7501 CTCTAGGCCC AAGAGACATT GCTTTATGGA TATAATTGTG TGTGTGTGTG AACATTGAGG
7561 AAAGGGAAAT AAAAAAATA CTTAGCCGCT TAAGGTTGTA CAGTTTCACT AATTGCTACT
7621 TTTAGTTGTG ATAAATGGC AGGTGCTTCA ACATTTATAT ATACAAAAAC TTCCCTGCTG
7681 GTGGTTCAAC TGTGAGAACT GGGGTAAGTG GGTGAGTTCT CTTTTCTGT CTCTGTCTCT
7741 GTCTCTCTCC TTCCATTCTT TCTTAAAGGA AATAAACATT GCAGCTGGGT TATAGCTCAT
7801 CAATATGGAA GTTACAGAAG TGAAAAAAGG CATTGCCTTG GTGGGTGGTG TTACCAGCTG
7861 ATTTTTGGTT GTCCTGCAAG GAGGTCTGGG GACTGGCTGC TCTGTCTCTG TCTGTATGAG
7921 TGAGGGAAAG CTGGGGAGCA GATTCCCTAA CCTTCAGCCT GGCCTGGTTC CTGAGTGAAC
7981 CCAGCCTCTC TGGTCCTAGT AGCTTTTTTCC AAACAGGAAT CTGAGTGGTG ACAGGGAACA
8041 AGTACCAGCC CATTGCTTAA GTGCCAGGGT TAGTGAGGGC AGGAAGCTGC CATAGCTGGG
8101 ATTAGTAGTT GTATTGGATG TAGGAAGTCC TATCCTGGGA CAGCTAATCC TTAATGCTTC
8161 ACTGGAGATT TTCAATGAGA AATTTATCCC ACGGCCCATC CTTTTGTCTC
8221 CAACAGCCAA GTATTTTCCA TTAGAGGAGA CTTCTGTAC ACTTGATGGA TGCTCATTCC
8281 AAGGTGACTT GGGGCAGTCA GTACAGACTT GGGATGACCT CTGACAGGCT AACCTCTCCC

Figure 23E:

```

8341 CAACAAGGGC CCTCTATGTT TGCTATGTAA TGTAATGTCA GACATTGTCA GGAGTGTCCG
8401 CAGCACAGCC TGCCCAGTGT GAGGGCTCTC ATAGGTTTCC CACTGTCTTA TCTACACAGG
8461 GATAACGAGG AGGTAAGCTG CAGTTCCAG TCTCACTTCA CAGAGGAAGA GATAACCCCA
8521 TCCCAGGTCA TGTAGCCAGC AGTGGAAAGA ATGAGGATTT GAACTCAGGT CTTCCAAGTC
8581 CCATTGATAG CATCTCCTCA CAAGTCCCTT GCCACCCTCA CGATGCCTTA GACACTTGCC
8641 TGCCCTTTAT ACTAAGGAGA TGCAGGTACA AGGGGTTTAC CCATGTAGCA GCTGAGGCAG
8701 CTGGGGATAG ATACCAGCAG CAGGCCTGAT GTCACCACCT TAACCTCCAGC ATCCCCAGTC
8761 TGTGTTTCTG GAGTGTGAAA ATCCCTACTT AACAAAGATT TGCAACAGTC CTTGGCTCTG
8821 TGACCCATAG CTGGAAACAG GATTCTCATT GATTGTGGA ACATGGTGGC AGCCAGCCAA
8881 AAAGAGGGTC TGCATACAGA AGACACGTGT GGCAAGGCCA CAGCAGACTC TGACTACCTT
8941 AGCTTACAGA ATTACAAGGT CATAATGTCC TCTGCTTTGG TCACCTCATG TTAAGGACAG
9001 GCCCTAATGA AGATGGGGCA GAAGACTGAA GGAATGGCCA ACCAATAACT GGCCCAACTT
9061 GAGACCCATC CTACAGGCAA GCATCAATTC CTGACACTAC TAATGATACT CTGTTATGCT
9121 TGCAGACAGA AGCCTAGCAT AACTATCCTC CGAGAGGTCC ACCCAGCAAC TGACTGAAAC
9181 AGAAAAAGAT ATCCACAGGC AAACAGTGGA TGGAGGTCAG GGACTATTAT GGGAGAGCTG
9241 TGGGAAGGAT TAAAAACCCT GAAGGGGATA GGAACCCAC AGGAAGACCA ACAGAGTCAA
9301 CTAAGAGACC TGTGGGAGCT CTCAGAGACT GAGCCACCAA CCAAAGAGCA TACACAGGCC
9361 GGTCCGAGGC ACCTGGCAGC TGTGAAGCAG ACATGCAGCT CAGTCTCCAT GTAGGTCTCTC
9421 CAATAAGCGG TAGCCTGACT GCAGTATCCA ATCCCCAACA GGGCTGCATA GTCTGGCCTC
9481 AGTGGGGGAG GATGCCCCTA ATCCTGCAGA GACTTGATGA GTGGAGAGCT ATCCAGGGGG
9541 AACCACCCCT CTCTGAGAAG GGAATGGGGA TGGGGGAGGG ACTCTGTGAA GAGGGGACAA
9601 GGACAAACAA GAACCTCAAA TAGGTCAGGC CCTAAAGGCT TGCTAAGTAG CAGTGGCCCA
9661 GCTCTGTCCT GTTCCTCAGC CCAAGGCTCA GCTCCACCT GTTTCTGTGT TTTTCTGGCT
9721 TTTTCATGGG CTAGGACTTG GTGACCAGTT CAAACAATGG GGCCTGTGGA AGACACAATA
9781 TACAAGACTA GGGACATTCC TGTTCTGCTG ACTATCCATA GCCTGATGTA GGTGGAAGGA
9841 CCAATCACT GGATTTCTAC CCTTGACAAA CCTTGACAGC TGAGGGCCTC TCAGAAACCT
9901 ATTTCTTCCA CTGAAAAATG AGACTCTCAA ATGAACGTCG TGACAATCAT CAGGCTTATT
9961 AAAGAGGTGT ATCTAACCTG AATGGCAAGC AGACAGCAGG CAAATGTCTG TATCAACCTC
10021 TAGGAAGGAC AAGAACTGCT CACTGCTGCC CCCCAGGAGG CCATTGTCTG AAACAGCTGC
10081 TCTCCTGCTG GTGCACAGGC CCTGCCTTCT CATTGCAGCC ACAGCCCTT CCTGTCTGAA
10141 CCTCCTGTCA GGTCACTGGG AAACAGATCA AGATGGAACA GGACAGCTCC TGATGGTAAA
10201 TAAAAAACAG TGGTCATGGC TATTCATAGG GGTTTATGCT TCTTCAGTCC AACTGTGAA
10261 GAGCTGTGGG CATGAACCAC AGTGTTCCAG GTAGAGTTGG GGTTCGAAA TTCACAGTGG
10321 GGTGAGCTCA GTAAATGTGA GCTGGAGGTC ACTCGTGAGA CACACAGTCC TGCTGCTTCT
10381 GTTCCCAATA TCCTGAGGAG ACGACACATC TACTTTGTTC AGAGGCCACA GTCTAGTTGA
10441 CCTGAGAGTT ACCAGTTTCT TATTTGTGTG TGTGTGTGTG TGTGTGTGTG TGTGTGTGTG
10501 TGTTGTTTCG GTGTGAGTGC AGGTGCACAT ATGATAGCGT ACACGTTGAG GTCAGAGGAT
10561 AACTATCAGG CGTTGTCCCC TCCTACTTTT CCTCGGACTC TGGAGAACA ACATAGGTCC
10621 TTATTCCAGG GGAGCAAGTC GCTGTGTGCT GACACATCTT GCTCACATAC ATTTTACCTA
10681 GACAATGGAG CCTCCATCAG AGTATTACTT TAGCTCCTCA CCGATGGCAA TGCACCACCT
10741 CTCTACCCAC ATAGGAGTTG GGTCTCCACA CCCCCCACA CCCCCTTAC CAAAACGTTT
10801 TCAGTTACTT TATCTGGTAA AGTTCATCAG AGAATGAAGC CAGTATTAAG AACATGGAAT
10861 CATTTGGGAA CCTGGATCTA GCAATACCC ACCCTAGATG GAGTTGCTGA GTTTTCACTT
10921 CAGATTATAA TTCCCCCCTA GCTTCTATGG TTTATTCTGA AACCAGGGGA ACTCGATTCC
10981 TCCCTTTGGA CCACAGACAT CCTGGCTTGT GAATTCACAT GTCATCTACT GCTAATCCAT
11041 TGGTAGTATG TGGCTCACAG AGACACACTA CAGTCATGGC CAATGTCAAG GTAGGACAGA
11101 TGTGAATCAT TCCCCAGTC CTGCTGTTTT CATGACTAAC CCTCCTCAGC ACAGTGACCA
11161 TGAACCTACT TTTCCCCTCC TTTTATTTTT AGAATTGCTG GAATTTTCTA TTTTGAGAAA
11221 TAATAGCCTT GGGCAGCATT AAACAAAATC ATCTAGAAA CTGGTTTAAA ATACAGATGG
11281 TTGAGTCAGT GAAAGAGTGA GGAATGTCTT TATTGGCCCC TCACAGAGGC TGGCTCACTC
11341 CAGCAGAGGT GGTGGAAGCT CTGGACAGG GGTGAGGTGC ATAGGAAAAG TNGTCTGGGA
11401 CACTGAGAAC CACAATTGAA CAAACAGAAC TGTTGGCTTT TTTTTTTTTT AATGAGTTCT
11461 CAAAAAATGA CTGGCTAGCT TAGGCAATA CTTCGAGCCA ACCCAACAGA ACATTCTTCC
11521 ATTGATTCAT TCTGGATCTT CTTTCTAGAC AATACTGAAC TGACCCCTTG TTGGCAGTCT
11581 CAAGTTTGAC AACATAGGGC TTTGAACTTG GCACAAGGTC CATCACTGTC ACCCAAGCAT
11641 CCTGGGTGAC CTTTGGGTG GAATATCTTG GCTAACCTTA GATATTTTCT TTGGAGTATC
11701 TTTAGAACAT CCAGGAAATA GGGCTTGATT CTCATCCTGG GACCACAATA TAAGTCACCC
11761 TAGAATCCCA GGAGATCGTG CAGAGAAACA AGGATCTCTC TCGTGTGCA CTTTCTTCAA

```


Figure 23F:

11821	AGCAGTGAGT	AGTGACTCCA	CTAAACTGAG	TTCCCATCTG	AGAGTCCACA	GGAGGCTTTG
11881	GGGCAAGAAG	CAGAGGGAAG	GCACTGTTTG	TGTTGGTAAA	GTTTTGACTC	TAACAAATTT
11941	GAAGACATAG	ATGACATTGT	GTCAGACTAA	CAACAACCTA	GACTCATGTG	GGTTCGTGTT
12001	AGGGATCAGA	TTTTATTTCAT	CAATGACTTG	TCTTAGTGTA	TAGAGAAAGG	CTTCCTACTG
12061	GAGTGTAGGC	TCAATAATGA	CAGAAGAGAT	AGCTATTTCC	CCTAGGGACT	GTGCTGCTCC
12121	AAGTTTGGTG	GAGAAAGGCA	GTGGGGAACC	TAGATGTGCT	CTCTGGGGAG	GGGGTCTGAA
12181	GCTGGCTTCA	TAGAAGGTGT	GAAGTTTTCG	TGAAACATCT	AAACAGAATT	ATAGCTTAGG
12241	AAAGTGAGCA	GGCAAGGCAG	GGAATGTGTT	GCATATGTAT	ATGTACATGA	ATATATTATG
12301	TTATAGATAC	ACACACATTT	GAACCTCATT	TGCAGATGAC	AGAAAATAGG	TTATTTTGCC
12361	TCTCTTAACT	GCTAAGCACA	ATGACTTCCA	GTTCCATCCA	TTTCCTGAAA	TGCCACAATT
12421	TCATTTTTCAT	TTGTGGCTGA	ATAAAATTCC	ATTGCAGACT	GGGCCCTACT	TCATCCACTC
12481	CTGAGGGCAG	GCATATCCCC	TGGCTCCATT	TCTTACCTAT	TGTGAAGAGA	AGTGCAACTG
12541	TCTTGTTGAA	AGGCAAGCGT	GAGAGAGGCA	GGCACTAATT	GTGGGTTTTT	GTTTCTTCTT
12601	CCTGCTATGA	CTCTCCATTT	GTCAGAACCA	AAGATCGATA	AAAGCCGCCA	CCATGAAAGC
12661	CATCTTAATC	CCATTTTAT	CTCTTCTGAT	TCCGTAAACC	CCGCAATCTG	CATTGCTCA
12721	GAGTGAGCCG	GAGCTGAAGC	TGGAAAGTGT	GGTGATTGTC	AGTCGTCATG	GTGTGCGTGC
12781	TCCAAACCAAG	GCCACGCAAC	TGATCGACGA	TGTCACCCCA	GACGCATGGC	CAACCTGGCC
12841	GGTAAACTG	GGTTGGCTGA	CACCGCGCGG	TGGTGAGCTA	ATCGCCTATG	TCGGACATTA
12901	CCAACGCCAG	CGTCTGGTAG	CCGACGGATT	GCTGGCGAAA	AAGGGCTGCC	CGCAGCTTGG
12961	TCAGGTCGCG	ATTATTGCTG	ATGTGACGCA	GCGTACCCGT	AAAACAGGCG	AAGCCTTCGC
13021	CGCCGGGCTG	GCACCTGACT	GTGCAATAAC	CGTACATACC	CAGGCAGATA	CGTCCAGTCC
13081	CGATCCGTTA	TTTAATCCTC	TAAAACTGG	CGTTTGCCAA	CTGGATAACG	CGAACGTGAC
13141	TGACGCGATC	CTCAGCAGGG	CAGGAGGGTC	AATTGCTGAC	TTTACCGGGC	ATCGGCAAAC
13201	GGCGTTTCGC	GAAGTGAAC	GGGTGCTTAA	TTTTCCGCAA	TCAAACCTGT	GCCTTAAACG
13261	TGAGAAACAG	GACGAAAGCT	GTTTCATTAA	GCAGGCATTA	CCATCGGAAC	TCAAGGTGAG
13321	CGCCGACAA	GTCTCATTA	CCGGTGCGGT	AAGCCTCGCA	TCAATGCTGA	CGGAGATATT
13381	TCTCCTGCA	CAAGCACAGG	GAATGCCGGA	GCCGGGGTGG	GGAAGGATCA	CCGATTACAA
13441	CCAGTGGAAC	ACCTTGCTAA	GTTTGCATAA	CGCGCAATTT	TATTTGCTAC	AACGCACGCC
13501	AGAGGTTGCC	CGCAGCCGCG	CCACCCCGTT	ATTAGATTG	ATCAAGACAG	CGTTGACGCC
13561	CCATCCACCG	CAAAAACAGG	CGTATGGTGT	GACATTACCC	ACTTCAGTGC	TGTTTATCGC
13621	CGGACACGAT	ACTAATCTGG	CAAATCTCGG	CGGCGCACTG	GAGCTCAACT	GGACGCTTCC
13681	CGGTCAGCCG	GATAACACGC	CGCCAGGTGG	TGAACTGGTG	TTTGAACGCT	GGCGTCGGCT
13741	AAGCGATAAC	AGCCAGTGGA	TTCAGGTTTC	GCTGGTCTTC	CAGACTTTAC	AGCAGATGCG
13801	TGATAAAACG	CCGCTGTCAT	TAAATACGCC	GCCCCGAGAG	GTGAAACTGA	CCCTGGCAGG
13861	ATGTGAAGAG	CGAAATGCGC	AGGGCATGTG	TTCGTTGGCA	GGTTTTACGC	AAATCGTGAA
13921	TGAAGCACGC	ATACCCGCTT	GCAGTTTGTA	AGGTACCCGG	GGATCACAAC	TTGCCCTCTG
13981	AAGAGGAAGA	ACAGAAGGAT	GCCACAATC	TCCTGCTGGC	TACTCTCCAG	TGGTTTTCATC
14041	TTACTTCTGA	TGGCATTTC	CTCTAGAAAG	TGCTACTATC	ATCCACACAT	TTCTACTGTA
14101	GACCACCCAA	AGGACCTTCC	CAAATTCTCT	TCCTCTCTGA	GTAGTCTCCA	CACCTGTTAC
14161	CACCATCCCA	GAATTAAAAT	CCTAACTGCA	CTCTGGCGTG	TGACTTGCCT	CAGTCCCTGC
14221	AATAAGAGTT	GTTGGCAGTG	CCAGGCGTGG	TGGCGCACGC	CTTTAATTCC	AGCACTTGGG
14281	AGGCAGAGGC	AGGCGGATTT	CTGAGTTCGA	GGCCAGCCTG	GTCTACAGAG	TGAGTTCCAG
14341	GACAGCCAGG	GCTATACAGA	GAAACCTGT	GTCGAAAAAC	CAAAAAAATA	AAAAAAGTT
14401	GTTGGCAGAG	TGTGGGTTAT	ATACCAGGTG	GAGATTTCAA	ATGAGTGGCT	GAAGCTGTAG
14461	CCAGAAGGAA	CTTAGAGGAT	AGCTCATAAC	TTAAAAAGAA	ATGTAGAGAG	TAGCAGAAAC
14521	ATTGAGAGAG	TGGGCACACA	GCCACTGTGT	GAATGTGGCA	GAACACAATC	CAGCCAGCTA
14581	TACATGCATA	AGTGTATATT	GGCGCCATCC	TGACTGATGA	GACACAGGAA	AACAGATAGA
14641	CGGGGTAGG	TGGCCATGGC	CTTTCCTGCC	TGCCTCTTCC	TAAGGGTCTAT	CTCAAGACCT
14701	TATGCTCTCT	TAAGTCTTCC	ATTGCTACTT	AGCTTCTAGA	TATCACCTCC	AGATTAGTCT
14761	CCTTGGGTAC	ATCAGTGATC	CTGGTGATAT	CCAGGGCTTC	CTGATTCCAT	CTTGTGCATA
14821	GAGGCTGCAA	CTAAAGAGGT	CTTCTTAATA	CTTCACACCC	TGATGCCAAA	AGGAAGACAC
14881	AGAAGTTCAC	AGAGGTGAAG	TGATTTCATG	AGGACATACA	GTGAGCAAGC	ATCAGGGTCC
14941	GGATTATCTG	ACTCTACTCT	AACTTTTATG	TAAATGTGCT	TTATGCCATT	AACACTGTCA
15001	TTCCTGTGCT	TCAGCTCTGG	GAGACTCCCA	AGCACTCTTA	GGCACAAGCC	ACAATTAAAGG
15061	GACTCTGACA	CTCTGCATTG	ATTAATTAGC	ATGGTGGTCT	CTATGTTTCC	AGATTCAATGA
15121	TTGTTTCACT	TTCCATATAG	GCTATGAAGG	GTGTGAGGAA	ATTTTTTGGG	GACAGAATTG
15181	GAGGCAATCC	ACCTCTCTCA	GGAAAGCCTCT	ATCTGGAAAA	GCTTACAAC	CAGGGACAGT
15241	AACTGTAGGC	CCAGTCCTTG	GTGTCCAAAA	TGGGTTTTAT	GTTTGAATC	TGCAAGCCT

Figure 23G:

15301 TCCATGTGCT CAAAGGTTTG AACATGGAGC CTCCTCCTGG TAACACTGTA TTGGAGGCTT
15361 TTGAGACTGG ATGCTCTTTG GTCCCATGTT TTGCTACATC ATCTGTCAAG ATATGACCCA
15421 GGCATGCTAC CAGCTACCAC AGACTATGCC TCTCCAGCTT TCATGTTCTC CCCACCATGA
15481 TAGACTTGTA TCTCCTAAAA ATGGAATCAA AGCAAACCTT TCCTGCATTA AGTTTTTTTT
15541 TTTCTGTAA GTGTTTGGTC ACAGGGACAA GAAACACTC AATACAGATA ATTAGTACCA
15601 GAGTTGAGGT TCATTGCTCT AGCAAGTTGG ATCAAATTTT TAGGGCTTTG GAACTGATTT
15661 ATAAGAGACA TGTAGAAGAG TCTGAAGCTG TGGGCTACAG AAGTGTCAAC AGTTTTTAAG
15721 AATAGTTTAA TACCCATGG GAATTGTGAA AATCAGAATG CTCACACAAA GGCAGACAGG
15781 AAAACGTGAG CATGTGGCGT GTGAGAGGGC ATAAGAAGGA ACCTAGGGGG AAATGAGCTA
15841 GAAGCCATTC GGCTACGTTA GGGAACGTGT GTGGCTGTGC TTGGCCCATG CCCTGGCAAT
15901 CTGAATGAGG CCAAATTTTA AAGGAGTGG CTAACCTGAT TGTCAGAGAA AATATCAAGA
15961 CAGACCACCA CTCAGGCTAT GCCGTGTTTG TGACCGACCA GCTACTCTTA GCCAGCTCTA
16021 TTGTGAAATT CCAGAGCAAT TATCAGAGCA TGAAGATACA TACAGTTTAG TGAAGTAAGG
16081 GGTGTGGGTC CCTAAGTGGG TGGTGCATAA ATCTATGTAG GTGATGCCTA AGTGACACTT
16141 GATAATCCAA AATATCAGCA ATGTGGAATG TCTTCCAAGG AGACCTGTAG ACACACATTT
16201 TAGAATTTTG CTCATGGCTG TAATAAATAG CTAGCTAGAA ATCATTTCCT GAAGAGGTTA
16261 GTCTGAGTTA CGGTTCCAGG GCAAACATTC AGTGATGGCA AGGAAGGCAT TGCAGTCAGG
16321 AGCCAAAGGT CAGCTGGTCA CATTGCATCA AGAGTAGAGA GTCAGAGTGT GAGTAGAAAG
16381 AGGATACAGG TTATAAAACC TCACTGTCCA CTCTCAGCAA TCCATTTTCT CCTAAAAGGC
16441 TTTACCTTCT AAAGATTTTA GTCTTCAAAA CCAGTACCAG TAGCCTGGGA ACAAAGTTG
16501 AAACAAATGA GCCTTTGTGG GGCATTCAC ACTTAAACA GGGCATCACC TAGGAGGAGC
16561 CCTGTGTGCA GTAGGAAGTG TGGCCTCTGT GTCAGGAATG CTCAGGCTAA TAAGGGGTCC
16621 TCTATCTGAG GGACCCTATG AAGATTCAAC AAGTAGTTGT GAGAATTCCC TGTAATGGA
16681 TGCTACCAAT TTGACATTTG TAGACCTGCT ATTGTGTGCT TCTTTATTGG GCTCTCCCAT
16741 CTCCCAACTT TCCAACCCAT ATTCCACATT AATCCCTTCC ACCACCATGC AACACTAGGT
16801 AGGAGAGAAG GAAGGTTAGA AGAGAAAGTG GGTATAGATC TATTTAGACT ACTTCCTGCT
16861 GATTAGGGGC AAGTCCAATC GTCATTGTCA GGATACCTCC AACCAGCAAC CAGCAAACCA
16921 GCAAATCAGA AACAGCAAAA GCAGCCAACA AGGCAGCACT AACCAGCAGG ATTGGGGTGG
16981 GTAGCGTGGG AGCAGTCACT ACTGGTCTTC TCATGGCTTT GGCATTAATA CTCTCTCAAG
17041 AAATTCGGTA ATTTTTCCTC CACCACCTGA AATTCGGTAA TTTTAAATGC AAATATCTA
17101 CAGCTGGCAA AAATCACATC TCTCCTAGAG CACAAGACAA ATCATAGTTA CTGGCTATTT
17161 GCAATCTGAA GCATCTCAAT ATCCACACCC TGGGATTAAA AAAAAACAT ATTCACATCA
17221 CATAACTGTT TTTTTTTTCC AATTTTTTAT TAGGTATTTT CTTTATTTAC ATTTCAAATG
17281 CTATCCCGAA AGTCCCCTAT ACCCTCCCAC CTCCTGCTC CCCTACACAC CCACTCCCAC
17341 TTTTGTACCC TGGAGTTCCC CGGTACTGGG GCATATAAAG TTGCAAGAC CAAGGGGCCT
17401 CTCTTCCCAG TGATGGCCGA CTAAGCCATC TTCTGCTACA TATGCAGATA GAGACACGAG
17461 CTCTGGGGGT ACTAGTTAGT TCATATTGTT GTTCCACCTA TAGGGTGGCA GACCCCTTCA
17521 GCTCCTTGGG TACTTTGTCT AGCTCCTCCA CTGGGGGCTC TGTGTTTTAT CTAATAGATG
17581 ACTGTGAGCA TCCACTTCTG TATTTGACAG GCACTGGCCT AGCGTCACAT GAGCCAGCTA
17641 TATCAGGGTC CTTTCAGCAA AACCTTGCTG GCATGTGCAA TAGTGTCTGC GTTTGGTGGT
17701 TGATTATGGG ATGGATCCAC TAGTCTTAGA GC

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☒ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.